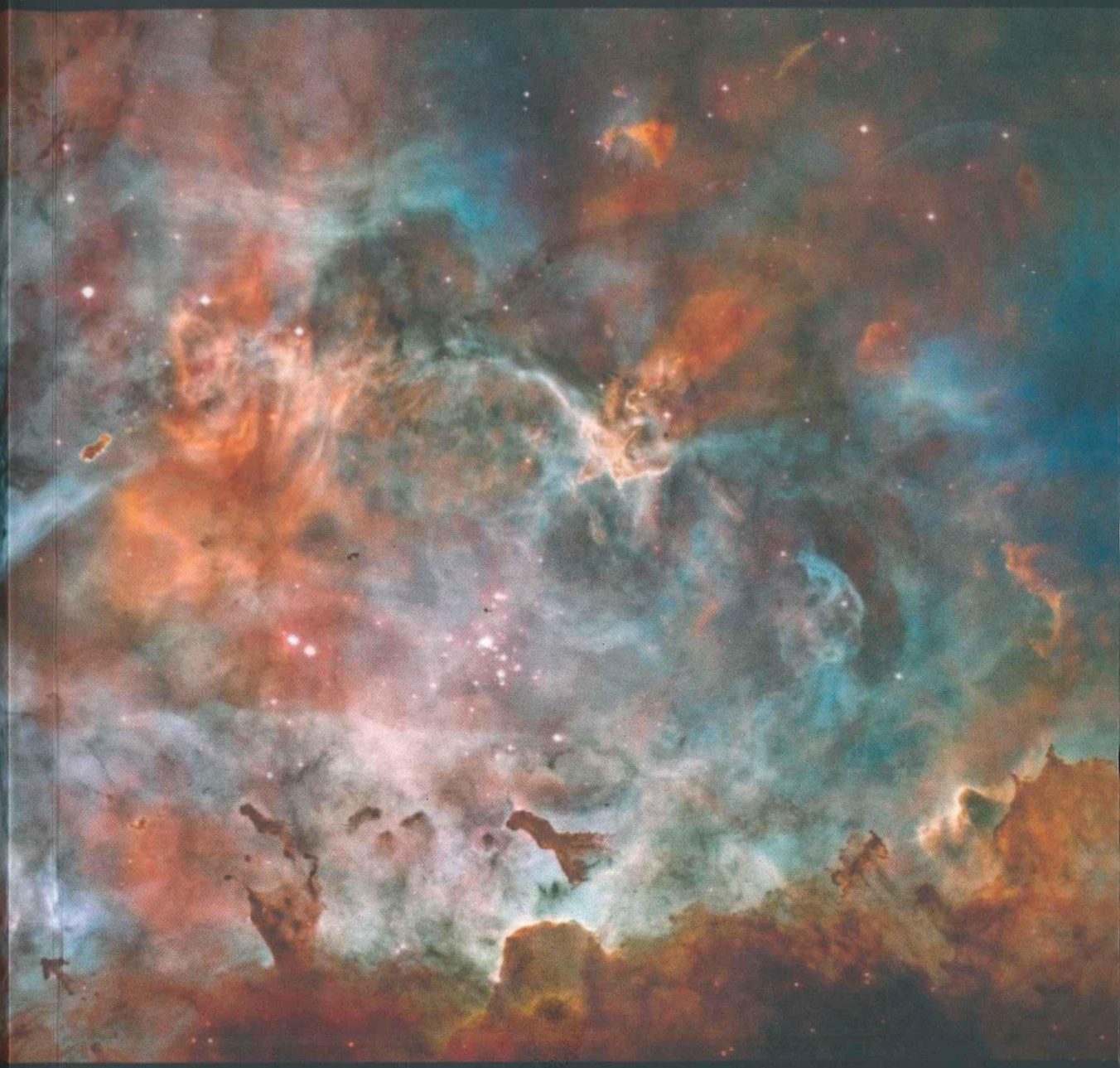
What is Physics?

A Defence of Classical Theism



Nigel Cundy

WHAT IS PHYSICS?
A Scientific Defence of Classical Theism

Nigel Cundy

To all those who asked me to describe how
my science fits in my views on religion.
I said that it would take a book to explain it.
Well here, at last, it is.

Contents

Chapter 1. Introduction 9

Chapter 2. Tools 39
2.1. Logic 39
2.2. Probability 45
2.3. The Scientific Method 55

Chapter 3. Metaphysics 63
3.1. Introduction 63
3.2. Assumptions and definitions 66

Chapter 4. Aristotelian Metaphysics 91
4.1. Introduction 92
4.2. The Premises of Aristotelian Metaphysics 95
4.3. Being and Matter 96
4.4. Mind and Life 104
4.5. Causality 114

Chapter 5. Aristotelian Physics 129
5.1. Overview 129
5.2. Failures 131

Chapter 6. The Emergence of the Modern World-View 135
6.1. Definition of Science 135
6.2. The Laws of Physics 138
6.3. The Emergence of Classical Physics 141
6.4. The Merton Calculators and Their Successors 145
6.5. The Experimental Tradition 152
6.6. The Emergence of Mechanism 165
6.7. Astronomy 178
6.8. Descartes 182
6.9. The Legacy of Descartes 189

Chapter 7. Classical Physics comes to Maturity 193
7.1. From Galileo to Newton 193
7.2. From Newton to Einstein 208
7.3. Space and Time: from Aristotle to Newton 223
7.4. The Formalisation of Mathematics 231
7.5. Special Relativity 232
7.6. Noether's Theorem 243
7.7. General Relativity 245

Chapter 8. The Philosophy of Classical physics 255
8.1. Positivism 255
8.2. Methodology 257

8.3. The Classical View of Space and Time 258
8.4. God 260
8.5. Causality 266
8.6. Is Enlightenment a Good Thing? 284
8.7. The Fate of Scholastic Metaphysics 287

Chapter 9. Statistical Approaches to Physics 293
9.1. Statistical Physics 293
9.2. The Second Law of Thermodynamics 297
9.3. Natural Selection 303

Chapter 10. Quantum mechanics and Field Theory 309
10.1. Introduction 309
10.2. Quantum Mechanics 315
10.3. Relativistic Quantum Mechanics 325
10.4. Quantum Field Theory 328
10.5. The Standard Model 349

Chapter 11. The Philosophical Implications of Quantum Field Theory 355
11.1. Form and The Physics of Composite Beings 355
11.2. Metaphysical Implications of Quantum Field Theory 362
11.3. Mechanism and Quantum Field Theory 367
11.4. Scholastic Metaphysics and Quantum Field Theory 369
11.5. Summary 374
11.6. Principle of Causality 379

Chapter 12. Modern Objections to A Classical theory of God 389
12.1. Introduction 389
12.2. Classical Presentations of the Argument For God 397
12.3. Hume 415
12.4. Kant 422
12.5. Summary 430

Chapter 13. A Classical Theory of God 433
13.1. Definition of God 433
13.2. The Attributes of God 450
13.3. The Mind of God 462
13.4. The Persons of God 487

Chapter 14. Ethics 499
14.1. Introduction 499
14.2. Types of Ethical Theories 507
14.3. Is an Objective Ethical Theory Possible? 520
14.4. Virtue, Duty, and Consequentialist Ethics 535
14.5. Natural Law Ethics 563
14.6. The Goodness of God and Man 597
14.7. God and Evil 606

Chapter 15. From Theism to Physics 629
15.1. Introduction 629
15.2. Premises 629
15.3. Quantum Field Theory 632
15.4. Fermion States 651
15.5. Bell's Theorem 654
15.6. Quantum Electrodynamics 658

15.7. The Classical Limit 670
15.8. The Standard Model of Particle Physics 673
15.9. Gravity 682
15.10. Recap 689
15.11. Fine Tuning 693
15.12. Summary 703

Chapter 16. Miracles 705
16.1. Introduction 705
16.2. Objections to the Miraculous 709
16.3. How to Assess the Evidence for Miracles 730
16.4. Summary 737

Chapter 17. Conclusion 739

Bibliography 743

Index 749

CHAPTER 1

Introduction

This book is primarily a quest for the truth.

Now, I fear that some readers might object that I am begging the question with my very first sentence: is there such a thing as *the* truth? In short, is truth subjective (i.e. varying from one individual to another) or objective (the same for everyone who reasons rightly and observes rightly, and not only everyone there is but everyone there could possibly be)? I will take it as a matter of definition that truth is objective. If it is not objective, then we need to find some other word for it.

So why might the idea that there is such a thing as 'subjective truth' have arisen? I want to mention three possible sources for this mistake. The first is in questions where there is ill-definition or contradictory definitions in play. 'This is the best motion picture ever' is a subjective statement, precisely because what makes the 'best motion picture' is ill-defined. People come in with different perspectives of what that might be, and thus although two people might say 'This is the best motion picture', they actually mean different things, even if they happen to agree on what the best movie actually is. So one of the criteria of a search for the truth is that everything must be defined, carefully and unambiguously, with any imprecision in the definition clearly acknowledged. If what made the best motion picture were defined in an unambiguous way (for example, if we define the best film as the one with the greatest percentage of running time given over to dialogue written by George Lucas – I doubt that many people would like that definition, but it is still a clear and unambiguous definition), then in principle all observers should agree, unless they make an error in observation or reasoning. I leave it to the imagination of the readers to work out what consequences this definition would have for the Oscars. But unless this is done, such statements are indeed subjective. But that is unimportant, because they are not truths, but only opinions.

A second possible source for this error is the idea that people from different perspectives might see different things. For example, one man might see a monochromatic light source as blue; another the same source as red. These testimonies could both be correct, if the observers were travelling at vastly different speeds. A more classic example is that one man may observe a disc as circular, and another as an ellipse, depending on what angle they are looking at it. However, though these cases may appear different, they are not really. The point is that a true statement of an observation should also provide a statement of the context in which it was made. The observer of the light should express his relative velocity to its source. The observer of the disc should state the angle at which he sees it. Once this is done, then the statements become objective, because each observer based on their measurement in their context can compute what every other observer in any other context would see. So this situation does not lead to the subjectivity of truth because the statements are, once expressed fully, equivalent.

A third possible source for this error is to include the effects of uncertainty of measurement or mistakes in reasoning. If not accounted for properly, these can lead to confusion. 'The wall is seven cubits high.' 'No, the wall is eight cubits high.' In reality, it is seven and a half, and each measurer has a different arm length, and are rounding their measurements in different ways. They should say something like, 'Because of the limited resolution of my equipment, the best I can say is that the wall is between six and a half and seven and a half

cubits high.' But this statement can be compared against reality, and found to be true, and with this error range both observers agree. Thus the uncertainty of measurement does not stand against the objectivity of truth. There is nothing wrong in saying 'I don't know' when there is still uncertainty. In fact, in these circumstances, it is wrong to say anything other than 'I don't know, yet.' Of course, the goodness of admitting ignorance only applies within the bounds of our uncertainty. Outside those bounds, we must say 'I know that it is not like that.' Mistakes in reasoning, on the other hand, lead to false statements. The existence of truth implies the existence of falsity (though statements can be classified in ways other than true or false), and this implies that it is possible to say false statements. The man who makes an error in reasoning is not demonstrating the subjectivity of truth, but only that he himself is wrong.

What about statements such as 'My foot is hurting?' This statement by itself, devoid of any context, is not something that can be called a truth. However, the context makes all the difference, because it clearly identifies 'My' with one particular person, and 'hurting' with that person's experience of pain, then 'My foot is hurting' translates into 'Joe Bloggs experienced pain from his foot at 3:12pm on June 15 2012,' and that is an objective statement, and thus capable of being true. The statement 'My foot is hurting', if spoken by Joe Bloggs at that time, can be combined with the context to give an objective statement (even if the evidence is not accessible to most observers).

Thus if we are searching for the truth, we need to use a methodology which a) promotes logical rigour; b) is precisely and tightly defined; c) deals with any uncertainties and imprecision in a rigorous and systematic manner; d) any statement given is complete, including all the relevant context; and e) allows us to transform between different perspectives observing the same quantity, or only deals with perspective-independent data. The scientific method is set up specifically to satisfy these conditions. Thus scientific observation is objective. There is, of course uncertainty, especially when constructing theories (we need not assume that every theory suggested by every scientist is correct; indeed those which contradict the experimental data are most definitely wrong), but the scientific method is directed towards finding the truth, or at least a part of the truth. We need not say that the scientific method leads us to the whole truth (indeed we cannot, since that statement itself cannot be proved using the scientific method), but it does aim at some of it.

Philosophy too promotes logical rigour, can be precisely and tightly defined, and at its best studies perspective-independent data. I am not certain that philosophers have traditionally correctly dealt with uncertainty (or even recognised the fundamental importance of experimental and theoretical uncertainty), but there is no obvious reason why they can't add some discussion of this into their systems, although obviously it would have to use different methods than are applied in the sciences. Thus philosophy too can be used in a search for the truth, with a broader range of applicability than the sciences (the sciences are natural philosophy and thus part of the overall philosophical endeavour). Similarly, the study of history can satisfy these criteria: there is a historical truth (what actually happened), and although we cannot know any of it precisely, we can still understand some of it accurately from the limited sources available to us. Therefore, while literary critics and the like may revel in the subjectivity of their fields, they are wholly wrong to extrapolate that to every field of study. There is a single truth, and I'm going to attempt to find as much of it as I can. So let us leave behind any childish notions of philosophical relativism, and move back to the hard objective world of scientific reality.

Truth, then, in its primary sense, means that a statement is in accordance with objective reality. It does not necessarily mean that the statement is perfectly precise, but a true statement is accurate. But is there such a thing as objective reality? This is by far the best explanation of our experiences. We know the inconsistencies of our dreams; we know that our waking life can surprise us in multiple ways which later turn out to be perfectly consistent. The success of physics, which assumes an objective reality, is strong evidence

that there is such a thing. And with physics ultimately giving rise to chemistry, biology, and consequently all other natural sciences, there is seemingly little room for a subjective a-reality to hide. Could it be tucked away in our minds? This is something that philosophers debate, but my feeling is no, since our minds are a part of that reality; ultimately arising from electro-chemical impulses in our brains (I leave open the question of whether this is a complete description of our minds, but it is certainly at the very least a significant part of it). If someone could study those impulses, and knew the key to interpreting them (I do not say that in practice this is possible for us, only that we cannot *a priori* rule it out), then our minds would not be private to ourselves.

Does this mean, then, that everything is contained within physics? That is the question which concerns me in this work, although the traditional answer is 'No'. But before we can answer this, we must ask ourselves what physics is. The difficultly with this question is that there are several different senses in which it can be understood. At one level, physics is the study of the motions and properties of matter, and once we have defined what is meant by 'motion', 'matter' and 'properties' I think that few people will dispute this definition. But we can also think of physics as an approach, combining painstakingly controlled experimental observations with mathematical modelling in order to understand matter, properties, and motion. Historically, this understanding is dubious, since there have been attempts to study physics without mathematics, and sometimes even without experiment, but as it applies to physics as it is studied today, few would disagree with this. Or we could think of physics as the current formulation of physical law in the diverse places it is applied, or the true and correct law that our current understanding is only an approximation to (albeit a very good approximation). But again, this formulation is not so controversial, except perhaps around the edges where our knowledge confronts our ignorance.

But there is a further sense of the question which is hotly disputed. What *is* physics, at the philosophical level? Describing the properties of matter is what physics does; the method is how it does it; the formulation of laws provide only a description, but a description of what? What is physical law, and how and why is it connected with reality? Some, of course, have regarded the universe as though it were a giant machine, and physics is just the manual to understand the workings of that machine, with each part a slave to pre-ordained and blind external laws. Others have regarded it as the grouping together of the natural tendencies of individual substances, and provides the connection between those tendencies with the form and nature of the substances. Others again would say that physics is a list of the possible ways in which particles can interact with each other, and provides a means of computing the probabilities by which those interactions can occur. Others would say that it is merely a construction of our minds, with no connection to reality. Many theists would say that physics is a description of how God sustains the universe in the normal course of events (excluding the miraculous). Each of these, and the many other answers that could be proposed, are an attempt to answer the various questions left unanswered by the mere formulation of physical law: Why those laws? Why does matter obey them? What is matter, or existence, or being in itself? How did the whole system start? Did it start? With the exception of the last of these questions, they cannot be answered by physics, simply because they lie beyond its scope. The science which does address these is the branch of philosophy called metaphysics, which literally means 'that which is beyond physics'.

Metaphysics is the foundation of physics. It is also the foundation of all other fields of philosophy (with logic being the tool of philosophy as mathematics is the appropriate tool for the physical sciences). Can we say anything about it? Some people dismiss metaphysics, saying that the questions beyond physics are unanswerable. But this will not do, partly because it is tantamount to giving up, but mostly because the statement that there is nothing knowable beyond physics is in itself a metaphysical hypothesis, one which requires justification. The problem is that those people who are repelled by metaphysics tend not to do this (quite naturally); in practice this means that they unthinkingly adopt various

metaphysical propositions (whatever they absorbed through their culture), only badly and without properly questioning them. If we are going to do things properly, we have to study metaphysics.

Metaphysics is, however, in some senses harder than physics, for one clear reason. Physics takes as its basis sensible objects; those things which we can see and touch. From these, and the process of deduction, careful comparison and falsification, we can deduce properties about other objects which we can't see and touch, and ultimately we derive physical law which describes how these pieces of matter interact with each other and move (again, I warn the reader that I haven't yet established what I mean by 'move' or shown that matter does 'move', or what I mean by 'physical law', or if there is such a thing). Metaphysics, done rightly, takes physical law as its starting point, and tries to deduce from that what lies beyond it. Thus theoretical physics is to metaphysics what experiment is to theoretical physics: first its inspiration, and then its testing ground. This means that it is very difficult to study metaphysics well. Most attempts at it are performed badly, which perhaps explains its poor reputation in the modern world.

It should be obvious from this discussion that metaphysics and physics are closely related. And this has been the case for almost all of human history, at least for as long as men have been systematically thinking about the deeper questions. Both subjects began with the ancient Greeks, and in particular, according to tradition, a man called Thales (although Thales himself was not from Greece, but what we now call Turkey, though he was still an ethnic and cultural Greek). He was not the first to think about the problems of how the universe works, but he was the founder of the way of thinking that eventually led to modern science. The Greek philosophers were inventive, and almost every idea in Western thought up to the twentieth century had some precursor in ancient Greece. We see people who viewed physics in terms of atoms (though they meant by this something more fundamental than what we call atoms today); people who considered physics geometrically; the earliest variants of both mechanism and theism; discussions of the possibility of vacuums or whether matter is continuous; the first questions of how we know and what we can know; whether there are universal kinds or just individuals with similar natures; and of course discussions about God. Among the Greeks were the first truly great geometers and mathematicians, and our modern mathematics has its roots partly in ancient Greece, and partly in ancient India. The Greeks were also the first great inventors and appliers of physics. In short, if we have a fundamental problem today, almost certainly some ancient Greek was thinking about it two and a half thousand years ago.

Of course, one of the reasons that the Greeks were long on speculation and short on conclusions was that they had not developed the required observational methods to sift through the possibilities. Equally, they were not taken to think things through with complete logical rigour. One man changed that; his name was Socrates. Socrates applied logic and clear thinking to these problems. We mostly know about Socrates through his disciple Plato, who became one of the most influential thinkers the world has yet seen. With Plato, the chain of thought begun two centuries earlier by Thales reached maturity. Why did Plato rise above his countrymen? There are two reasons: firstly, his thought was genuinely superior: more logically rigorous, based on clearer foundations, accepting the best ideas from his predecessors and rejecting a large proportion (though not, unfortunately, all) of the worst, and explaining more of the world; secondly because later cultural shifts made the Platonic school particularly appealing. There is a strong affinity to monotheism in the works of Plato and his disciples. Obviously Plato himself was a polytheist; however there was a clear distinction between the religion of the great philosophers and the religion of the Greek priests. Monotheism was a natural conclusion of Plato's system of thought; though Plato himself didn't bring it out (perhaps because he didn't see it; perhaps because the fate of Socrates warned him against going too far against the culture of his time), this became more obvious as Plato's thought was systematised over the subsequent centuries.

Plato's philosophy was therefore adopted by the early Christians as they began to work out their theology, and deployed by them in their intellectual battle against the Pagans and the Manicheans. The victory of Christianity ensured the victory of Plato.

Plato was, perhaps, the first great metaphysicist; he was also the last for a long time who was both systematic in his philosophical thought, and yet was relatively unconcerned about natural philosophy, what we now call the physical sciences. His clear and logical thought led him to many great achievements, but his physics was limited, and therefore his metaphysics was limited.

But Plato had a disciple, a foreigner, a Macedonian, named Aristotle. Aristotle was not as highly regarded as Plato in the ancient world; his surviving writings are considerably drier and harder: Plato was one of the few great philosophers who was also a great stylist and writer. Aristotle was also not different enough from Plato that he was clearly distinguished from his old teacher. Compared to the great variety of Greek thought, Plato and Aristotle certainly both seem very similar. Thus Aristotle was treated in the classical period as just one member, albeit an important one, of Plato's school. However, their differences were also important, and here I want to briefly mention two of them. For both Plato and Aristotle, a goal of intellectual endeavour was to understand forms, which is what different objects of the same type have in common. For Plato, these forms dwelt in an intellectual and abstract world, and were distinct from the material objects we observe. Thus to gain a true understanding, one has to pierce through the material world to reach the higher spiritual world beyond it. For, Aristotle, however, the form is a part of the material objects, and is best understood through detailed study of those objects, and how they change and react to stimuli: knowledge comes first of all through the senses and observation. The second difference is related: Plato borrowed from Pythagoras the idea that the fundamental layer of reality, which Plato interpreted as his collection of forms, was mathematical and geometrical in nature. Plato and his students thus strongly emphasised the study of geometry, believing that it led to an alternative and superior way of understanding form than starting from material objects. Aristotle, however, was interested in sensible (in its original meaning of being apprehensible to the senses) objects: rocks, plants, animals and so on. These are clearly not abstract mathematical objects, but concrete realities. Aristotle therefore promoted the idea that study of the natural world was (largely) divorced from a study of geometry; his physics was entirely non-mathematical while that of Plato's school was rooted in mathematics.

Like Plato, Aristotle's philosophy was naturally compatible with monotheism. However, unlike Plato, Aristotle was also highly interested in natural philosophy. He was the first true empirical investigator of science (though obviously he did not develop his methods anything like as well as the modern scientists). He was the first to treat logic systematically. He was the first to study living organisms in depth, and wrote one of the first major treatises on the biological sciences. He wrote on metaphysics and ethics, rhetoric and poetry; but, most importantly for my purposes, the first major systematic and complete works on the subject of physics. In Aristotle, physics and metaphysics were merged together, and interconnected, and thus began the long relationship between the two. This was the first revolution of physics, and its birth: the two hundred or so years of Greek speculation which culminated with Aristotle and the first systematised treatment of physics which worked well enough to be used practically. Aristotle's physics was a combination of his own metaphysics and studies, and also drew various premises from his predecessors among the Greek philosophers.

Of course, Aristotle's own work on physics contained many rough edges. It was gradually improved and polished by later writers over the course of centuries. Most importantly, Aristotle demanded that there were two types of natural locomotion: linear locomotion towards or away from the centre of the universe (i.e. the centre of the earth), and circular locomotion around that centre. Terrestrial objects moved linearly; the heavenly objects moved in circles. His physics actually works far better than most of us today might suspect,

but problems were found, most importantly with his account of the heavens. The locomotion of the planets was carefully observed, and found not to fit with Aristotle's theory. However, his work was improved, and Ptolemy in his *Algamest* seemingly perfected the system, combining Aristotle's metaphysical principles with a model of the solar system which was consistent with the best observations of the time. The work of Plato and Aristotle continued to be developed during the Greek, Roman and Byzantine eras; it merged into the Neo-Platonic school, and was well set to unleash the full potential of science upon the world. I should here mention John Philoponus, who lived shortly before the Arab conquest of Egypt, and might have launched the Galilean Revolution nine hundred years early, had not history intervened.

But history did intervene. The emergence of the Roman Empire, with its disciplined government and militarism stalled and slowed the development of Greek thought, concentrating more on politics and the practical rather than the natural sciences; the collapse of the Western Empire under barbarian invasions destroyed all education in Latin Europe; the emergence and conquests of Islam stalled thought in Persia and the Eastern Roman Empire, and eventually led to the decline of philosophy in the East. Initially the peace and huge empire of early Islam acted as a melting pot. There was a great hunger for learning in ancient Persia, which early Islam inherited, and as the Caliphate's armies brought back to Baghdad the intellectual as well as physical treasures of the lands they conquered and raided, from Greece and North Africa to India, the influx of new ideas stimulated new thoughts. In particular, it was during this time that Aristotle came to be truly separated from Plato, and his thought came to prominence. It is important neither to overstate nor understate the contribution of the mathematicians, philosophers and scientists who lived in the early centuries of the Islamic Caliphate (some of them Jew, some Christian, some Zoroastrian, and some Muslim); they should not be understated, since they did more than just preserve the ancient knowledge, but systematised it; separated the good from the bad; made some important advances, most importantly in mathematics, their monotheistic interpretation of Aristotle, or in their emphasis of the experimental method in studies of optics; and merged together the learning of almost all of the civilised world at that time (with China being the most prominent exception), particularly bringing in important insights from ancient Indian philosophy and mathematics which otherwise would not have reached Europe. However, neither should their contribution be overstated: they did not make the key breakthroughs that marked the transition from ancient Greek and Indian science into modern science, and after a few centuries Islamic philosophy began its long decline.

Following its initial conquests when Islam first burst out of Arabia, the Islamic world found itself in possession of various major centres of learning in Mesopotamia, Persia and North Africa, and the newly conquered scholars, still permitted to practice their original religions (but not to convert anyone from Islam) played an important role in the education and bureaucracy of the new Islamic Empire. Thus was born the Mu'tazilite school of Islam, which tried to merge the Arabian religion with the rationalistic philosophy of the lands they had conquered. It was this school which produced the achievements of Islamic philosophy which would go on to inspire the Europeans. However, the Arabs themselves had no intellectual tradition, and suspicion grew that the philosophers were supplanting the authority of the Koran; the Asharite school of Islam emerged in response to the Mu'tazilites. While Christianity emphasises the *logos* of God, which can be translated as word or rational nature, the Koran emphasises the supremacy of God's arbitrary will, meaning that according to what is now orthodox Islamic theology, philosophy cannot be used to understand either God or the natural world, since both are seen as being beyond human rationality. Several of the natural philosophers of this period clearly stepped outside the bounds of what was allowed by Islamic theology. There was a real danger that rationalistic science might challenge the Islamic religion. Thus the study of natural philosophy, after a brief and glorious flowering, was suppressed. Philosophy was only useful to refute Islam's enemies or interpret

legal texts. The great Mu'tazilite scholars (who would later be so honoured in medieval Europe) were dismissed as heretics, and their works and influence destroyed. Islamic philosophy thus became limited and declined, and made no more significant contributions to the story of the development of contemporary science.

It was, however, in Western Europe, after its emergence from the dark ages, that the story resumes. It took a long time for Western Europe to recover from the fall of Rome. First there was the invasion of the Germanic and Hispanic tribes that brought down Rome, then the Scandinavians from the North and the Arabs from the South continually raided the remaining embers of Latin Christendom. None of the conquerors had an intellectual tradition; none favoured systematic education, and the political chaos and constant fear of brutal and bloodthirsty raiders from both North and South made a general academic education seem like an unaffordable luxury. Gradually, however, the continent was re-Christianised, and the raiding parties expelled; the Scandinavians, Germans and Slavs converted and took a leading role in European civilisation, while the Saracens were partially subdued by various counter-attacks and reconquests. By the ninth century, political and religious stability was restored, and education began again. Of course, at the start there was little to build on baring the few scraps of Plato preserved in the monasteries, but beginning from the eleventh century, either from trade or conquest of Arab and Byzantine colonies in Spain and Italy, the classics of Greek and Mu'tazilite thought were translated into Latin and introduced into the emerging universities. In particular, Aristotle was re-introduced to Western Europe. Among Aristotle's works was his physics, and with this, in twelfth century Europe, scientific progress resumed. Slowly, at first: in particular it would be some time before the importance of the experimental method was fully accepted.

Aristotle's physics was based on his views of causality; in particular his idea of final causality that matter had a natural tendency towards certain ends and not others. Things fell down when released because the element of earth had a natural tendency towards downwards motion. He did not link terrestrial physics with geometry. In late medieval Europe, that was changed. Various scholars tried to describe Aristotle's physics geometrically, and they were successful. This led to a new concept: if locomotion could be reduced to geometry, and geometry to various axioms, then that suggested that matter moved as it did not because of an internal cause propelling it, but because there was some external rule compelling it. This law was not variable but fixed, just as the laws of geometry were fixed. Thus the philosophy of mechanism was born: that physics can be reduced to matter in locomotion (with locomotion the only form of motion), matter cannot be created or destroyed, but must obey fixed and deterministic laws. Equally, the ancient philosophy of nominalism was restored: that there are no definitive kinds in nature, but only accidental similarities, with everything bound by the same laws. In this way, the second revolution in physics and metaphysics began, and once again it was driven by metaphysics. The metaphysics changed, and then people, realising that Aristotle's physics was inadequate in several details, began trying to find the laws that their new philosophy said had to be there. Flaws were found in Aristotle's terrestrial physics, and in particular the idea began to emerge that forces led to acceleration rather than velocity, and the precursors to the concept of inertia started to emerge.

It was, of course, roughly contemporary with this that the experimental method began to be practised again, though, for a time, the geometrical natural philosophers and the experimental natural philosophers remained separate. There was also a renewed interest in Astronomy, and a revolutionary theory was suggested that it was the earth that moved around the sun, rather than the sun revolving around the earth. The scientific revolution of Medieval Europe was almost still-born; the black death decimated Europe's universities, and the Renaissance, which turned attention away from Aristotle with his albeit flawed focus on natural philosophy towards writers such as Cicero who focused more on politics and human society, nearly undid all the earlier progress. Fortunately, the second revolution was

saved, by the emergence at about the same time of three geniuses trained in the best of the Medieval scholarship: Galileo, who merged and combined the geometrical, astronomical and experimental strands of early modern physics (and showed that their combination could go far further than any of them on their own); Kepler, who perfected the Astronomical models; and Descartes, who provided several major advances in mathematics required by the new physics, and was among the first to present a truly fleshed out philosophy based on the mechanistic principles. The second revolution in physics culminated with Isaac Newton, who discovered the long sought after laws of motion, developing the remaining mathematics he needed at the same time. This triumph seemingly confirmed the mechanistic view of the universe. Aristotle's physics and metaphysics were swept away.

It was also at about this time that philosophy started to split into numerous different schools, for the first time since the victory of the Plato/Aristotle consensus at the beginning of the Christian period. Descartes might have been a genius, but his philosophy had numerous difficulties. However, the mechanistic school continued and developed, removing or ignoring these difficulties, and became the dominant philosophical school accepted by the physicists themselves. The other two schools of philosophy were the empiricists, and the nominalists. The empiricists, of whom Hume was the dominant member, restricted human knowledge to only what can be observed. In particular, the empiricists distrusted all metaphysics, mechanism as much as Aristotle's, and argued that anything conceivable could happen. This, of course, represents an overthrow of the study of physics itself, because physical theory, being induced from experiment, is among those things that Hume would cast into his fire. Theory is any essential part of physics; the goal of physics is to come up with a good theoretical understanding, even though this is tested against experiment. The nominalists, of whom Kant is the best example, steered a middle ground between Hume and the mechanists, by dividing reality into two parts: the external world which we cannot truly know (as suggested by Hume), and the internal world of the mind, in which Newton's Laws held firm. Newton thus was not governing the external world, but only our understanding of it. From Kant also came the idealists, who (to over-simplify a little) decided that the external world of physical objects was superfluous, and omitted it from their metaphysics. Thus the link between physics and philosophy was partially broken. The mechanists remained (who were most closely aligned with Newtonian physics), but were not the only school of philosophy, and many of the others thought in ways which diverged from the most natural interpretation of physics. But all these schools of philosophy were (and are, since they continue, though much developed, to this day) in agreement that Aristotle's old metaphysics was wrong, and in particular the distinguishing features of his system: formal and final causality. These philosophers are grouped together under the name *Enlightenment*. It is said to be a period of great philosophical reawakening; the most productive period in philosophy since ancient Greece, perhaps even surpassing that. The difference, of course, is that while the philosophy of the Greeks eventually moved to a consensus around the Platonic school (which included, as far as the Greeks were concerned, Aristotle and others), enlightenment philosophy seems to be more divided than ever. But, I suppose, it is still early days. It took more than half a millennia for Plato's ultimate triumph over his ancient rivals; the modern philosophies haven't yet had that long.

There are thus three different major streams within modern Western thought, two rational and one irrational[1]. The irrational stream (though perhaps anti-rational would be a better name) I call post-modernism, and it is characterised by the idea that there is no objective truth, or the truth is unknowable, or (worse) that we can (and do) mould the truth according to our own wills, desires and prejudices. Thus, for example, the human

[1]Of course, my 'three' here is somewhat a simplification; firstly because there may well be other minor thought systems, and secondly because many people take various ideas from one stream, and other ideas from another, and try to merge them together. Equally, within each general stream there are many different currents.

condition is seen as what we ourselves believe it to be, and as malleable as our own beliefs. If a man chooses to self-identify as a teapot, then reality is warped accordingly and he is, indeed, a teapot, and anyone who denies it is decried as a bigot, hater or teapotaphobe. And of course, for the post-modern mind, such claims are justified: if there is no objective standard to judge him against, then the only reason to deny him would be blind prejudice or a desire to upset him. To the rational mind, though, such a conclusion is risible, and such name-calling a symptom of the post-modern's underlying irrationality. The rational mind would respect the integrity of anyone who reaches a conclusion on the basis of logic and evidence, even if they believe that the evidence was misinterpreted or incomplete, or the logic based on false (but nonetheless plausible) premises. There is no point in insulting someone who disagrees with you, or accusing them of bigotry when they are as or perhaps even more rational than you are; one instead tries to correct them, or accept correction yourself. There is an objective reality, which we can know well enough to make a judgement, and the reality is that he is a man, and if his beliefs do not conform to that reality, then it is those beliefs which need changing and are an evil. At best such a man needs correction; at worst he is insane. Another symptom of post-modern thought is the invention of new categories that overlap different real categories, and the attempt to shoe-horn people into those categories. For example instead of people who happen to be black (among other qualities), it (in its most extreme form) constructs the community of black people, and treats it as though it were a single entity; instead of treating individuals as a mixture of good and bad characteristics, it tries to express everything in terms of rivalries between different groups, with both the rivalries and groups existing only in the post-modern mind's imagination. Such over-generalisation and simplification only gets in the way of uncovering the individual causes of the individual examples of strife that exist in our world, which are far more complex than the post-modern narrative will allow. Such thought, of course, reduces to prejudice, sentimentalism, and the rule of the will (most importantly the will and prejudices of whoever holds temporal authority, whether judicial, political, teaching or economic), and unfortunately it is dominant in much of what passes for ethical, political and social thought in public secular culture.

The two rational streams of thought, on the other hand, believe in an objective reality, and that mankind should subject himself to truth rather than try to be its master. These are the modern secular school, which arose from the mechanistic philosophy of the Enlightenment (most of which evolved into first deism and then atheism), and the religious school (which remains Christian or some related belief), which retains (sometimes consciously, sometimes unconsciously) a residue of pre-enlightenment scholasticism. These schools are distinguished not by their commitments to reason or to the importance of empirical investigation, for both hold such things in exceptionally high regard, but merely that they proceed from very different and mutually contradictory premises. Of course, many members of the secular stream believe the religious stream to be irrational, and not a few members of the religious stream believe atheism to be irrational. The reason for this is that when evaluating the other school of thought, such thinkers unconsciously bring in premises from their own understanding. Such premises are contrary to the axioms of the other school, so naturally when one works through the details various contradictions arise. But these contradictions mostly do not come from within the stream, but from the assumptions used to analyse the stream; if judged on their own terms and accepting their own axioms, each stream is logically self-consistent.

Certainly, adherents from one stream would say that their own premises are drawn from the results of natural science and are thus impeachable; the problem is, so do the adherents of the other stream. Before we interpret natural science, we have to make various assumptions; the premises supposedly extracted from natural science are often drawn from these assumptions. For example, it is believed by both schools that matter behaves in a regular way. But are these regularities external to the matter, something imposed on

the matter; or are they internal to the matter, and arise as a consequence of the matter's fundamental nature; or are they due to some compromise position between the two (stating that the regularities are a matter of the internal nature of the substance, but that the nature was imposed on the substance from outside, and that the regularities are often - or always depending on the school of thought - actualised by something external to the substance)? Science, which merely provides a description of these regularities, cannot answer. However, philosophically, there is a big difference: the ideas of external laws of physics to which matter is bound is characteristic of some streams of the mechanistic approach (and a denial that the question makes sense to other streams of mechanism); while the idea of internal natures is characteristic of the Aristotelian understanding, and the compromise position leads to Aquinas' version of classical theism.

Physics itself, of course, continued to develop throughout the eighteenth and nineteenth centuries, oblivious to all these philosophical debates. Newton's triumph gradually brought more and more phenomena under its sway. It seemed that nothing could stop it, and that gradually all of physics would be united under the sway of Newton. Astronomy; optics; the phenomena of heat, entropy and equilibrium; the statistical treatment of large numbers of particles; the physics of fluids and gases; electricity and magnetism: one by one these fell to the geometrical and mechanistic physics. Newton's own laws were expressed in other ways: the principle of least action, and Hamilton's equations, each of which gave new insights into mechanical physics. At the same time, there were hugely important advances in mathematics, most notably (from the perspective of a physicist): complex analysis and the theory of complex numbers; the development of group theory and the studies of symmetry; of topology; and non-Euclidean geometry, most particularly that of Riemann who (along with Gauss and the other early pioneers in this field) caused something of a sensation in the nineteenth century: Euclid, seemingly the most secure of all Greek thinkers, had also been overthrown. However, aside from complex numbers, these were not yet needed by the physicists.

But Newtonian physics finally fell. Two problems emerged. Newton's first triumph was to explain Astronomy, but it was Astronomy which would prove to cast doubt on his theory: minor deviations in the orbits of the planets (particularly Mercury) differed from what was predicted from Newton's theory of gravity. The second problem emerged from the combination of electromagnetism and optics. Maxwell's equations suggest that either the speed of light should be the same for each observer or the strength of the electrical interaction between two particles depended on how fast they were travelling. Both of these hypotheses violate Galilean Relativity, one of the foundations of Newton's theory, and both are strongly counter-intuitive. When we measure the relative velocity between us and a truck, it depends on how fast we ourselves are travelling: we will get different results if we are standing still (relative to the earth), walking or on a bicycle. Why should light be any different? Experiments soon showed, however, that it was. The speed of light in a vacuum is constant for all observers no matter how fast they are travelling.

This was, of course, the spur for the third revolution in physics, the theory of relativity in both its special and general forms. Although Einstein didn't realise it when he first set out, the key idea followed from the mathematical work of the previous century on symmetry and non-Euclidean geometry. Einstein (or more accurately Minkowski, based on Einstein's initial work, though Einstein was quicker than most to recognise the importance of what Minkowski had done) realised that instead of thinking about the geometry of space, we should be thinking of space and time as a single entity. But this geometry was not Euclidean, but the hyperbolic geometry first discovered by Gauss. Changes in velocity could be represented by global symmetry transformations of this four dimensional coordinate system. A global transformation is when we rotate the whole of our system of parametrising space (or space-time in this case). Take a piece of paper, and draw x and y axes on it, crossing at the origin, so the x axis is pointing towards you and the y axis is perpendicular to it. Now

rotate the paper around the point where the axes cross. That's a global transformation, in this case a rotation. A local transformation, however, is a little bit different. In a global transformation, the same transformation is used for each point on the coordinate system. In a local transformation, we change the coordinates differently at each point.

Once Einstein finished with different velocities, he started also thinking about what happens when you start accelerating. This, it turns out, corresponds to a local transformation of the coordinate system. The mathematics of local transformations are very different to global transformations, particularly when we turn to calculus, the key mathematical advance which allowed Newton to write down his equations of motion. We need to add to the differential operator (the key component of Newton's equations) a connection which depends on the reference frame and even the particular path the particle is travelling along. Fortunately for Einstein the mathematics he needed had been worked out by Riemann, though it took him several years to wrap his head around the philosophy of what he was doing. He was overthrowing the notion of absolute space, a key part of Newton's philosophical framework. However, Einstein solved his problems, published his general theory of relativity, and, as a bonus, found that he had a theory of gravity as well. In Newton's equation, Gravity was a force that had to be added to the dynamics of his theory. For Einstein, Gravity was caused by the fact that space-time is not Euclidean. Effectively, if the principle of inertia states that objects continue in a straight line unless acted on by an external force, then Einstein more carefully defined what is meant by a 'straight line', and used that to explain motion in a gravitational field. Einstein's theory of gravity works, of course, better than Newton's[2], and so far no experiment has contradicted it.

The problem that motivated the two theories of relativity came from physics, but Einstein's solution was motivated as much by philosophy (and Mach should be credited here for inspiring Einstein's line of thought). The mechanistic world view that motivated Newton needs to be modified, but those modifications are relatively minor, and mostly relate to the nature of space and time; the nature of matter is left unchanged. But the main advance of Galileo, Descartes and their medieval predecessors, that space and time should be thought of geometrically, remains. It's just a different geometry. But as far as the philosophers are concerned, Einstein's advances didn't change very much. Indeed, they can be thought of as the perfection of the insights that led to the mechanistic revolution.

So what of Philosophy? By this time, the strongest position was dominated by mechanism, but including various insights from Kant and Hume; a position known as modernism. In particular, we should here think about God. The traditional position of Christianity (as well as Judaism and some schools of Islam) is theism, which is the belief that God not only created but also actively sustains the universe. Every motion in the universe is an act of God; though God acts in regular ways and uses the natural properties of matter as secondary causes. This position is fully compatible with the Greek philosophy of Plato and Aristotle; indeed, as Thomas Aquinas showed, the premises of Aristotelian philosophy taken to their logical limits demand it. However, the mechanistic view of the universe allows a different possibility. Here the action of sustaining the universe is delegated to fixed laws of motion. This is not necessarily incompatible with theism – they could just be a description of God's sustaining of the universe – but most of those thinkers treated them as having an independent existence to God. They were at first motivated by theology – a divine law-maker naturally suggests the existence of laws – but God was reduced to setting the system in motion, establishing the laws, and perhaps occasionally giving the universe a little push. As more and more science was understood in terms of the mechanistic laws, the need for God to push became less and less. This is the position known as deism, which

[2]Newton's theory of Gravity is equivalent to Einstein's if the difference from Minkowski space-time is very small; in the solar system, only mercury travels in enough curvature of space-time for there to be an observable difference between the two theories. Thus Einstein not only thought of a theory of gravity, but also explained in terms of it why Newton's theory works so well.

in its strongest form states that God is an idle observer of the universe apart from the act of creation; the weaker forms allow for intervention from God (for example, in the Biblical miracles), but the universe is still granted an largely independent existence. Why have God in there? One argument for God was weakened when it was noted that Newton's laws had no natural starting point: they seemed to suggest that the universe continued eternally into the past in a steady and stable state. The only real argument for God remaining was the argument from design, particularly the biological argument where we see that the diversity of natural life forms, and their own internal workings, is ordered incredibly well, far more so than the most complex of human machines. That surely has to be the work of an intelligence? Not according to Charles Darwin. Thus what was seen as the last good reason for believing in God (according to the academic philosophers) disappeared. The deistic consensus of the mid-nineteenth century became an atheistic consensus at the level of academic philosophy, with perhaps only a few Kantians and die-hard churchmen dissenting. By the early twentieth century this consensus was overwhelming, and though it has diminished a little since then, it is still prevalent.

Of course, that was just at the academic level. At the popular level, God was still very much in vogue; though eventually the doubts of academia started to seep through into popular culture, starting with the university educated politicians, civil servants and teachers, and from them more generally into society. The Roman Catholic Church's response was to try to revive its old philosophy of Aristotelian Thomism as an alternative to the mechanistic world view (on which the new atheistic consensus depended); orthodox Protestants (the evangelicals) generally retreated into their own world paying little heed to secular philosophy, but concentrating on their own methods of scholarship; liberal protestants accepted the premises of the new philosophy, but tried to deny its conclusions, and generally remain deists of one variety or another.

It should be noted that this consensus was based on the mechanistic philosophy which seemed to be inconsistent with theism. Is it really inconsistent with theism? We still have to ask the question of what the laws of nature are. Many early scientists regarded them as a description of God's upholding of the universe, and the consensus that they should have an independent force apart from God emerged later: God was removed, but nothing put in His place. The laws became the most fundamental entity: but this is troubling. Theologians have no problem in making God fundamental, since in classical theism the fundamental entity must possess various attributes which uniquely belong to God. But the laws of physics, and the universe as a whole, lack those properties; therefore they cannot be fundamental; yet the question of what lay beyond the laws of physics was not one which nineteenth century atheistic philosophy could adequately answer. However, neither do the laws demand a personal God in the same way that the older Greek philosophy does. But mechanism itself, with the full weight of Newton and later Einstein behind it, seemed unquestionable.

But there has been a fourth revolution in physics, with roots in the early years of the twentieth century, but reaching its climax in the 1970s, and this time it caught the philosophers completely by surprise. It was not philosophy-driven, like the first three revolutions, but driven wholly by experimental physics. Nobody saw quantum mechanics or quantum field theory (as the mature form of quantum physics is known) coming. And unlike Einstein's advance, the effect of the new physics on philosophy ought to have been dramatic. Instead, for the first time since Aristotle, physics and mainstream philosophy have become divorced. This is not a problem for physics, but it is a problem for the philosophers.

Quantum mechanics began with a number of observations. The first was the observation that accelerating electric charges radiate electromagnetic radiation, and electrons in a classical orbit around a particle nucleus are accelerating electric charges. They ought, if classical mechanics is right, to continually emit radiation, losing energy as they do so, until they collapse into the nucleus. A stronger element of this problem is that the emission

spectrum of atoms (of which hydrogen, being the easiest atom to solve, is most important) is not continuous, but comes in various lines of definite energy. This was shown to imply that the electrons around the hydrogen atom had to have certain fixed angular momenta, integer multiples of a parameter called $\hbar$. In a classical picture, this makes no sense. A second issue which spurred the development of quantum mechanics was the observation that electromagnetic radiation, known for some time to be a wave, also showed in certain circumstances particle like behaviour; with each individual lump or quanta of radiation having an energy of $\hbar\omega$, where ω is the angular frequency of the light wave, and $\hbar$ the same constant. Curious that the same quantity is appearing in two seemingly unrelated places.

The third observation was that electrons, one of the fundamental building blocks of matter, showed diffraction effects, which are characteristics of waves. But these arrived in a strange way. You could actually observe individual electrons hitting the detector. Individually, the pattern seemed to be random, but when a large number of electrons passed through the experimental apparatus, the overall intensity took on the classic interference pattern characteristic of light or water waves. In water waves, this is easy to understand - you have two different signals coming into contact and combining with each other – the interference pattern arises because of the interactions of large numbers of particles. The same pattern in an electromagnetic field was assumed to be from the same cause - interference between different wave-fronts: classical optics is based on this assumption. But when we have a single electron arriving in that same pattern – what is it interfering with? Also note the phrase I threw out a few sentences ago: 'seemed to be random'. By this I mean that it is unpredictable. You could fire a single electron out of the same apparatus in the same set-up and get a different result. Well, could there be something in the set-up that differs from case to case? Too small for us to measure, but big enough to make a difference to the final outcome? The answer is 'No', or at least it is not the most important reason for the result. Other unpredictable physical processes were also observed: spontaneous radioactive decay being one of them. What this means is that at a fundamental level, physics is indeterminate, by which I mean that it is in principle as well as in practice unpredictable.

This should get mechanists worried. Determinism is a key part of the mechanical philosophy. Furthermore, there are issues with the forms of causality proposed by enlightenment philosophers. Enlightenment causality (with the exception of Hume's confusion between causes and correlation) is mostly what I call *event causality*. Changes in motion are caused by events, say two particles happening to collide, or by forces. But in spontaneous radioactive decay, we have something changing without any event causing it, nor any force driving it. The idea behind causality – that every change in motion has a cause – seems to be very much undermined. Thirdly, the mechanistic philosophy is centred around the idea that the fundamental principle of physics concerns particles in motion: understand the movements of particles, and you in principle understand everything. Waves, the other conveyor of energy and momenta apart from particles, could be reduced to local motions in a fluid, such as water waves or sound waves. But now we have a confusion, or duality, between particles and waves. The fundamental building blocks of nature seem to be neither one thing nor another.

The mechanism that inspired Newton and was refined by Maxwell is, in short, dead. This might not be a disaster – maybe it can be adapted or modified. But, while that first generation of quantum physicists worried about the philosophical issues, they laid them to one side, and continued to do what physicists do best: modelling the observed world mathematically, and coming up with a mathematical theory that fits the observations. The theory of quantum mechanics was developed in different ways by several different people: Schrödinger, Dirac, Heisenberg, Pauli, and others; and the various different models were shown to be equivalent: all these great minds had converged on the same answer, albeit expressed in different ways. The basic idea was that instead of particles being the fundamental

building block of nature, the key to physics is something called a wavefunction, which represents a probability amplitude for a single particle. A probability amplitude is something you can use according to a fixed rule to compute a probability; that probability describes the unpredictability of quantum observations in a systematic way. In other words, we can't identify where a particle is going to be found when we go looking for it (this is where the indeterminacy comes in), but we can ascribe a certain probability where it could be seen in each place, and this probability can be computed from the wavefunction. The wavefunction in quantum mechanics evolves deterministically according to a fixed equation the non-relativistic Schrödinger equation being used most often: it's wrong (its non-relativistic, so can't get the physics right at high momenta corresponding, in non-quantum physics, to velocities close to the speed of light), but it is easy to solve in the most important circumstances and it is also good enough for most purposes; the indeterminacy appears in the actual measurement process. The evolution described by the Schrödinger equation leads naturally to the interference effects; mathematically there are distinct similarities between quantum mechanics and classical optics (though, of course, there are also important differences). Measurements are associated with the second key element of the formalism of quantum mechanics, known as operators.

There are a few aspects of Schrödinger quantum mechanics which are a bit ugly – we should be worried that it is non-relativistic – but it works exceptionally well. The experimental evidence poured in. New advances came in in diverse fields. Finally we could understand the atom, chemistry, molecules and solids. Staggering new types of solid were predicted and then discovered – we owe our modern electronics and materials industries to those early pioneers of quantum mechanics.

The next step was to merge quantum mechanics with special relativity. It was already known in principle how this ought to be done. The problem was that in practice it didn't quite work out. There were issues with states of negative energy – even potentially infinite negative energy. This was good in some senses – it lead to the prediction of the existence of anti-matter, which was soon verified experimentally – but it was also troubling, because matter naturally falls down into lower energy states, so why isn't the ground state of the universe at an energy of $-\infty$? More serious, however, was the observation that relativistic quantum mechanics in some circumstances implied negative probabilities, which is a mathematical nonsense. Equally, the situation arose when you could have one particle going into an experiment, and, if it had enough energy, three particles would come out. This is perhaps to be expected: Einstein's famous equation $E = mc^2$ means that energy and mass are in principle convertible, so there is no good reason why some kinetic energy can't be converted into mass, i.e. a new particle. It doesn't actually happen in Einstein's classical (in the sense of non-quantum) theory of relativity. But it does happen in the quantum theory of relativity. But this is bad because according to the formulation of quantum mechanics, the wavefunction describes a single particle: but here we have new particles being created from the first.

Of course, the philosophers should start getting even more worried. In the original formulation of mechanism, the fundamental objects of matter are meant to be indestructible, neither can they be created. Theory, and later experiment confirmed it, showed matter being created and destroyed.

What these anomalies show is that quantum mechanics isn't the right answer. It is along the right lines, and provided the key to getting the correct answer, but its not quite there. In the late 1930s, Dirac started to work on a theory which had the creation and destruction of matter at its heart: quantum field theory. He encountered problems with the mathematical consistency, mainly that various results gave infinite values when trying to calculate them, which he failed to solve. It was next generation of physicists who solved these problems, via a method known as renormalization, and in the late 1940s and early 1950s, a successful relativistic quantum field theory was developed which accounted for the

electromagnetic interaction. The process of renormalization was suggested by Bethe, and Feynman, Tomonaga, Schwinger, and Dyson were the key minds who turned Dirac's idea and Bethe's proposal into a workable theory. And quantum electrodynamics really did work, with experimental verification to an incredible accuracy.

The strong and weak nuclear forces were a bit more troublesome. The basic idea was first formulated by the then young Chinese scientist Chen Ning Yang (in partnership with Mills) in the mid-1950s, but it took some twenty years for the consistency of Yang-Mills theory to be shown (principally by 't Hooft and Veltman) and the full theory to be developed. In the mean time, experimental particle physics kept discovering new particles. The electron, neutrino, proton, and neutron, and their anti-particles, where simple enough to deal with; but then numerous other particles started making themselves known, and a new unifying recipe was needed, just as the atomic picture had unified the diverse chemical elements. This was provided by the quark model, in which the proton, and neutron where no longer fundamental, but composite particles[3] made up of three other objects given the unfortunate name of quarks. Other members of the zoo of particles were found to be constructed from either two or three quarks.

The combination of the Yang-Mills theory and the quark model provided the solution to the nuclear forces. The weak nuclear force was the first to be solved, with the final theory, which incorporates the electromagnetic interaction, written down by Weinberg, Glashow and Salem. The theory of the strong nuclear force came a little later, by Gross, Politzer and Wilczek. The electroweak model and quantum chromodynamics are also quantum field theories, and are again successfully tested by experiment to a very high precision. Together they form the standard model of particle physics, established since the early 1970s. Experimental verification of the standard model has continued, with the discovery of the Higgs Boson, first confirmed in experiment in 2012, completing the set. So far there is only one experimental result which contradicts the standard model: the discovery that neutrinos are not massless; but there are several minimal extensions to the standard model which can explain this and I believe that it will be resolved through a combination of theory and experiment in the near future. The problem is only a matter of deciding which of the proposed extensions is right.

Finally, I should mention Kenneth Wilson who, in the early 1970s and building on the work of others, put the renormalization method onto a much stronger mathematical footing. Before this work, the need to renormalize quantum field theory was a fundamental weakness of it; after Wilson's work it was seen as natural given the premises behind quantum field theory. I also note that classical mechanics arises naturally as a limit of quantum field theory, when we take the parameter $\hbar$ to zero. Classical mechanics thus emerges as a first approximation to the true theory of quantum field theory, valid when lengths multiplied by changes in momenta are large compared to $\hbar$, but with corrections of increasing importance as the length scale or momentum scale decreases. $\hbar$ is very small; when performing a classical experiment it is so small that the effects of it being non-zero are lost in the experimental uncertainty. Thus quantum physics not only passes the test of experiment at a microscopic level, it also correctly predicts the results of all the experiments that were used to support Newton's physics, while contradicting the mechanistic philosophical basis behind Newton's original formulation of his laws.

The study of general relativity also developed during the latter part of the twentieth century. It was shown that the universe hasn't extended infinitely into the past, as the early twentieth century physical and philosophical consensus believed. It is of a finite age; there was a point where it began. This again has been confirmed by experiment; a standard model of cosmology has emerged, alongside the standard model of particle physics, based around the idea of a 'big bang' origin to the universe.

[3]Note that I am using the word *particle* following a convention. I do not mean the same thing by the word as when discussing a classical particle. These are quantum, not classical, objects.

Is the standard model of particle physics (plus neutrino masses) the final, complete, theory of nature? No. The standard model represents the merger of quantum mechanics and special relativity; indeed the basic principles behind the standard model represent the only consistent way of doing this. There remains the issue of merging general relativity into the picture, something which has stumped some of the best minds in physics since the 1950s. There are, of course, many models proposed, with very grand things said about some of them by their advocates, but without experimental guidance we physicists are shooting in the dark. Other theories have been proposed to formally merge quantum chromodynamics and the electro-weak theory (they sit somewhat separately in the standard model), but here there is no good theoretical or experimental reason why we should do so. However, it is fully expected that the final theory will be a quantum field theory. Merging curved space time into non-quantum physics made no major changes to the underlying metaphysical picture. I see no reason why merging it into quantum physics should alter the ideas behind quantum field theory. We cannot be sure, of course, but we can be very confident that quantum field theory is the fundamental framework of physics, and therefore the right place for metaphysics and the interaction between physics and philosophy to begin. The question is not so much whether it is a quantum field theory, but which quantum field theory.

All of these advances in twentieth and twenty-first century physics were experiment-driven, and not philosophy-driven. There is no philosophical system which can claim the credit for having inspired them.

This means, of course, that there is no reason to expect that any philosopher before the 1950s came up with the correct metaphysical theory, and if they constructed their ideas using the best physics available to them (as a metaphysicist should) then that philosophy is almost certainly going to be wrong because the best available physics at that time was fundamentally wrong. Of course, if the older philosopher developed their system without reference to physics (as Aristotle did: his physics was derived from his metaphysics plus some additional assumptions, partly observational and partly drawn from his predecessors; his metaphysics was developed independently based on thoughts concerning the nature of change), or contrary to physics, then by chance they might have reached the right answer. But all of them needed to be tested against modern physics. Pre-quantum field theory philosophy might not be wrong, but (baring studies into logic) none of it should be assumed to be correct.

So how did philosophy respond to quantum field theory, or the earlier quantum mechanics? In general, not very well. Mainstream philosophy has, as far as I can tell, remained wedded to the ideas of either Descartes, Hume or Kant. One new response in the twentieth century was the rise of post-modernism, which focused on the subjectivity of knowledge, but that is a dead-end: all physics, whether quantum physics or classical physics or even Greek physics insists on the objectivity of knowledge. From what I can see, there is now a shift back to the modernism of the early twentieth century. Of course, many philosophers have studied quantum mechanics, and there are numerous ideas and hypothesis here with no real settled consensus. There are, however, two problems here: outside the particular sub-field of the philosophy of quantum physics, few philosophers seem to pay much heed to the issues, or even understand them; and secondly the discussion is almost universally of quantum mechanics, not quantum field theory. There are important differences between the two, and it is not clear that a philosophy of quantum field theory would have a quantum mechanical analogue, so trying to extrapolate from a philosophy of quantum mechanics to a philosophy of field theory might just be a waste of time. Philosophers should skip over quantum mechanics, and focus on field theory. A few individuals do work on the philosophy of quantum field theory, and they are to be commended. The dominant picture among them seems to be the Everett multiple worlds picture. Basically the idea is that the indeterminacy of quantum field theory is resolved because at each event, the universe branches off into two. We live on one of these branches, and only have access to that branch, so to us the

universe seems to be indeterminate; but from the perspective of the whole multi-verse, it remains deterministic. In short, there is another version of me in a parallel universe, exactly the same up to the point when I wrote 'branches' in the previous sentence, and he wrote 'streams', and from that point onwards our two alternative universes will diverge. This idea has become a staple of modern science fiction, but it has always left me uneasy; for one thing, it is not in principle experimentally verifiable. For another, it begs the important question 'Why on earth is the universe set up in this crazy way?' That the universe is at its heart indeterminate is by far a more natural solution[4].

All of philosophy, most importantly ethics as the most practically important branch, ultimately depends on metaphysics. We can only access metaphysics by considering what fundamental physical theories are consistent with the metaphysical world-view, and eliminating those world-views which are inconsistent with physics. But contemporary metaphysics is divorced from contemporary physics. We need a new approach to metaphysics, and based on it, a new approach to all philosophy.

A new approach. Or perhaps an old one.

The idea which lead to this book began with an observation that the main reason, and in my view the only good reason, that Aristotelian metaphysics was abandoned (and the search for an enlightenment metaphysics begun) was because of a perceived incompatibility with the physics that culminated with Newton. But why does that matter? Newton was wrong at a fundamental level. So if Aristotle's metaphysics, as developed by the Islamic and Medieval Christian scholars, is indeed inconsistent with Newton's physics, that doesn't say anything about whether or not it is true. Indeed, the very aspects of Newton physics which were used to undermine the Greek metaphysics were the very parts of it which were turned around in quantum theory. This doesn't mean that Aristotle was right – he certainly wasn't: he got a lot of things wrong. But it does mean that we ought to repeat the initial work of the enlightenment but this time based on the new paradigm of quantum field theory, and compare Aristotle's philosophy against our modern understanding of physics, to see if it is a possibility. Aristotle's physics was, of course, based on his metaphysics, but also on various other assumptions about nature. That Aristotle got his physics wrong simply means that those other assumptions were wrong – but we already knew that. The early modern philosophers raised other objections against Aristotle (such as that some of his concepts were incomprehensible), so we have to consider those as well. I am not, of course, saying that we should investigate Aristotle because it is good to imitate him. I will feel free to modify or discard any aspect of his philosophy which is inconsistent with quantum field theory, and add anything that field theory requires; and interpret things in the light of modern knowledge rather than the ancient categories. My quest is for the truth, not to vindicate a man who died over 2300 years ago; I use him only to give me ideas which I then need to test.

However, what I want to argue in the course of this volume is that Aristotle should be given another serious study by contemporary philosophers. His metaphysics fits surprisingly well with quantum field theory. Not perfectly well, of course: we need to make some changes and alter it in certain respects. But well enough that it is a serious contender for the physics based philosophy that we are searching for.

This, of course, would explain why modern philosophy has had such a hard time with quantum mechanics. Enlightenment philosophy was founded on the denial of two aspects of Aristotle's system: formal and final causality[5]. Yet both of these concepts arise naturally

[4]It is interesting that the original reason why a determinate universe was insisted upon was theological: it allowed God to predict and thus have foreknowledge of the future. In practice, this is both bad theology and bad physics.

[5]Don't misunderstand: there are several incompatible definitions of *final causality*, and the one accepted in the early modern period is not what Aristotle or his most important medieval commentator meant by the term. Unless you are trained in Aristotle's thought, it is likely that you have a different definition for the term than the one I use here.

in field theory. What about causality? Quantum physics is incompatible with the *event causality* proposed in the enlightenment based on the mechanical world-view, (and in the modern world the only form of causality that most people consider); but Aristotle's causality is different, what I call *substance causality*, and that is again compatible with quantum physics. Indeed, field theory demands a form of substance causality.

But if Aristotle's metaphysics (or some minor modification of it) is correct, and the mechanistic metaphysics wrong, then there are consequences, in particular with respect to the question of God. The consensus among philosophers has been that the mechanical philosophy most naturally leads to some form of either deism or atheism (I am not convinced that this is correct; but it is certainly a widely held-view, and one strongly argued). This is responsible for the situation in our modern academic elite in the Western world, where most people, even churchmen, are practical atheists or deists. It was, however, shown in the high middle ages that Aristotle's metaphysics naturally implies some form of theism; and to my knowledge this has never seriously been disputed by those who understand the philosophy. Modern attempts to address those old arguments usually either misstate them, bring in enlightenment presuppositions which have no place in this particular discussion, or, more usually, both. Similarly ethics is ultimately built on premises drawn from metaphysics. The modern view of ethics is rather diverse, being a mixture of utilitarianism, relativism, voluntarism, and human rights based approaches (which different ethicists having a different mixture of these core elements). But Aristotle's metaphysics necessitates a virtue-based ethical system, which, if his metaphysical premises are true, is far more rigorously derived than any of the modern efforts, and which disagrees with modern morality in a number of crucial ways.

I should, before continuing, address a misconception. The enlightenment is known as the 'age of reason', with the implication that everything that came before it wasn't. That is false. There have, unquestionably, been dark ages in mankind's scholarship (or more usually lack of it), but the high middle ages in Western Europe – from the eleventh century until the black death hit the continent – was not one of them. The scholars of that time were devoted to logic and reason, far more than the enlightenment philosophers. They were later ridiculed for it during the renaissance. Compared to this period, the enlightenment was actually an age of irrationality – though only, of course, in comparison to the greatest medievals, and only if rationality and irrationality are to be defined in terms of appreciation of and use of clear logical argumentation. The medievals did lack something crucial, which the enlightenment philosophers had: the experimental method. If the premises of the Medieval scholars were correct, it was, without experimental testing, only by a stroke of incredible good luck. But we should not doubt that their conclusions rightly follow from those premises just because they lived eight hundred years ago. We should never think that our ancestors were idiots in comparison to us; they weren't, and most of them would view our time (baring our technological expertise and much of the mathematical sciences) as the age of illogical idiocy (certainly the quality of political discourse on both sides has not been so poor for a very long time). We should not fall victim to chronological snobbery.

So now we move to the modern science-religion debate (which has, of course, been raging since the nineteenth century). Many atheists, of course, claim that their viewpoint is scientific. In practice, this argument breaks down to three steps:

(1) The listing of a particular set of metaphysical premises A.
(2) A claim that the premises A can be drawn from the findings of modern science.
(3) A sequence from reasoning from A to the conclusion either that if the premises are correct then there is no God, or that there is no intellectual need for God, or that God's existence is sufficiently unlikely that we need not bother about it.

Obviously, the order in which these steps are done might vary from person to person; historically step 1 came first, and provided the impetus that began the scientific revolution of the sixteenth to eighteenth centuries; step 2 followed based on the success of modern

science, and its conclusions fed back to fine tune the premises and give a stronger formulation of step 1, and step 3 followed during the later part of the nineteenth century, and reached its climax in the early twentieth century. Neither do I wish to claim that every modern 'scientific' atheist works from the same or even fully consistent premises; there are certainly great similarities between all these thinkers, but not identity in every detail. Step 3 of this process, as I have listed its, is logical deduction. It is therefore in principle secure if done rigorously enough. It has been challenged by various branches of liberal theology, who accept the premises A but reject the conclusion that there is no God, but to my mind (as a theist) the atheists here make a vastly superior case to the theologians: if the premises are correct, then the conclusion that there is no God follows. The focus then turns to step 2 of the process; this is an argument from induction, and therefore more dangerous than the deductive argument of step 3. Showing that the premises are consistent with the findings of modern science is one thing; showing that they (up to minor variations) are the only consistent set of premises is much harder. We should also bear in mind that the process of validating the premises was largely performed in the era of Newtonian (and pre-Newtonian) physics; physics has moved on since then, and we should not assume that the arguments made in the sixteenth to eighteenth centuries remain valid in the light of modern physics.

However, a theist rejects the conclusion that there is no God because they reject the premises A, and the contention that they necessarily follow from the findings of modern science. This is sometimes difficult for the atheist to grasp; the premises A are so entwined in the modern world-view that it is sometimes difficult for the modern thinker to avoid sneaking them back in through the back door. Thus the theist has a similar three steps in their argument:

(1) The listing of a particular set of metaphysical premises T.
(2) A justification of the premises T either drawn from the findings of modern science or from basic pre-scientific observations about what that the world is regular implies.
(3) A sequence from reasoning from T to the conclusion that God exists, and that various attributes of God logically follow from the definition and the premises T.

Once again, I will add the caveat that not every theist will accept precisely the same set of premises T, but that such variations are sufficiently minor for the purpose of this discussion that we can neglect them. By pre-scientific I do not mean historically or that it is drawn contrary to or without reference to the findings of modern science, but instead with reference to the order of knowledge: these are assumptions that must be accepted prior to any scientific investigation. If science is the study of physical regularities, then it must make the prior assumption that there are regularities to study, and the standard justification for T is that they follow from this observation alone. If the universe is regular and predictable in any way, then the premises T follow.

The listing of the premises was first made by the ancient Greeks, and the conclusion of step 3 first drawn at about the same time, although the case was considerably strengthened in the Islamic period, in medieval Europe, and to the present day, with the premises also refined over time. I do not believe that this case has ever been significantly challenged; the arguments proposed by early enlightenment thinkers such as Hume and Kant (and their successors to the modern day) tried to transpose the arguments from the metaphysical world-view defined by T into the world-view defined by A; this transposition naturally led to contradictions, which they pounced on. But they did not address the arguments as originally delivered, nor show that they failed within the framework established by the premises T. A few modern scholars have tried to argue against the conclusion that T implies God, but on my reading such attempts have failed. The overwhelming majority of philosophers and philosophical theologians today who accept T (who, of course, constitute a significant minority of philosophers and even theologians) would also state that these premises imply

the existence of God. I will present my own set of arguments to justify step 3 in chapters 12, 13 and 14.

This brings us to step 2. This is the point where the theist argument was most vigorously challenged; the Renaissance period spawned a great deal of scepticism of the medieval world-view, and the early experimental evidence against Aristotle's physics implied to many people that there must also be a problem with the metaphysics T, which, coupled with the observation that the rival principles A were spawning a new physics that met (as it seemed at the time) the challenge of experiment led people to abandon T in favour of A. But this conclusion is a little hasty. Firstly, it was not shown that the premises T were inconsistent with the experiments of the time, only an early application of them which brought in additional assumptions as well: the experiments showed those additional assumptions to be wrong, but not the premises that were used to justify God's existence. Could T, or perhaps a minor modification of T adjusted to take into account modern physics, also be compatible with the new experimental findings? In all the excitement of the time, this question was not answered, and though those who accepted T later (indeed several centuries later) made the case that it could, they were ignored or shouted down. But again, all this discussion is irrelevant: the comparison was made between T and Newtonian (and pre-Newtonian) physics, but physics has moved on. The comparison should be between T and modern physics.

Thus the science/religion debate should be centred around the question: which set of premises, A or T, is most consistent with the findings of modern physics? Of course, one could also propose various other sets of premises: I do not wish to claim that these two world-views are exhaustive, merely that, as the foundations of modern atheism and of theism, they are the most relevant to our contemporary Western society. In this book, I intend to argue that, although classical physics perhaps favoured A (though not conclusively), modern physics is fully consistent with the premises T, which I outline in chapters 3 and 4. I do not go as far as ruling out A entirely, although I believe this to be a natural implication of my work; nor do I discuss any possible alternatives to A or T. Thus this work should not be seen as a proof of God's existence; it should rather be seen as a demonstration that atheism has no right to label itself uniquely as scientific, and that instead the premises behind classical theism are fully compatible with modern science, and perhaps considerably more so than the premises behind modern atheism.

My intention in this work is to present the positive case for classical theism, steps 1-3 of the argument above, as well as I am capable of doing (whether my intention will be fulfilled is for the reader to judge). To establish something as true, however, requires both to demonstrate that it is self-consistent and consistent with the observed evidence, *and* that *all* its rivals are not. I will not disprove all of classical theism's rivals. That is partly because this book is long enough already, and partly because to discuss every philosophical system in sufficient detail to disprove them would require some superhuman effort of which I am neither qualified nor capable. I will therefore limit my discussion of other world-views to only a few others which historically have interacted with theism. My main focus will be on a few versions of the mechanical world view, which replaced theism in the early modern period, and remained dominant, albeit in a modified form, into the twentieth century.

This book is not a refutation of Hume, Kant, Descartes or any other major author who opposes my own thesis: it is not intended to be. I only address authors such as this to show that objections raised from their own world-view have no force in mine. There are also many perspectives which I do not discuss at all: from Western philosophy idealism, existentialism, the works of Wittgenstein, Frege, Nietzsche, Marx and so on, and nothing from Persian, Indian, Chinese, or Manichean thought, nor the numerous other non-Western philosophical systems. This is not because I want to claim that such writers or their systems are unimportant or incoherent, but only because I lack the space to discuss them in detail and, from what I know of them, their thoughts do not significantly interact with my own

world-view or offer strong objections against it. Of course, this means that this project is incomplete: a full discussion would have to engage with such authors, and the many more that could be cited, but that is better done by a philosopher rather than a physicist. As a physicist, my task is to show that my own system is consistent with contemporary physics; I leave it to the proponents of those writers to present their cases, and their direct opponents to try to show them inconsistent. The charge that I am offering a false dichotomy between theism and mechanism would be true if my purposes were different; my goal, however, is to show that classical theism is consistent with the world as we observe it, not that only classical theism is so consistent – a much harder task.

I focus on the mechanical world-view as an opponent of theism for two main reasons: firstly, its importance in that it is unconsciously accepted by many in our modern society. It is what is (usually implicitly rather than directly) taught in our schools, and still forms the basis of several major contemporary schools of philosophy; secondly because of its historical role. Classical philosophy was rejected because it was believed that the mechanical philosophy was a better match with the emerging physics, and it has not yet been seriously considered beyond that time with the misrepresentations of the early modern writers treated as though they were both correct and decisive, while (as those few who do study classical philosophy have long lamented) they are neither. Obviously, as physics has developed further, other world-views have been proposed to challenge mechanism, but they have been assumed to be superior to classical philosophy on the grounds that if mechanism is better than the classics, and the modern approaches are better than mechanism, then the modern approaches are better than the classical systems. By showing that the first leg of this argument is false, by demonstrating that classical philosophy does in reality a far better job in describing contemporary physics than the mechanistic approach, then I hope to undermine the assumption that the other, more modern, approaches to philosophy must be better than the classical system, with that assertion made without comparing the predictive power of the various systems in detail. Of course, it may be (as far as the arguments I develop here) that some alternative philosophical system does as well as I do here, or perhaps even better; that, however, would have to be the topic of a different book.

I point out that this work is primarily a work on philosophy (albeit one that only a physicist could write), but I am not a philosopher (in the sense of the modern academic discipline), but a physicist. I have tried to familiarise myself as much as I can with the works of the most well regarded philosophers in preparation for writing this; but obviously I am not an expert, and don't have an expert's breadth of knowledge. It may be that the devastating counter argument I am unaware of is contained within the writings of a minor philosopher; it may be that I misunderstood something important in those which I have studied. I have tried to answer what seem to me to be the most serious objections raised by modern philosophy as I go along, but naturally the reader should judge what I write critically. There will be objections I overlooked, either because I encountered them but considered them sufficiently weak that I had no need to discuss them, or because I have neither encountered them nor thought of them myself. Of course, that I have not voiced an objection does not mean that it cannot be answered; but nonetheless, I would rather this work were regarded as a first, preliminary, foray rather than the definite case for my ideas.

I should also mention my own bias before I start, so that the reader can judge my writings accordingly. As a Christian, I have a tendency to prefer certain lines of argumentation. Of course, being a Christian is not incompatible with having a high regard for the truth, or demanding that one's faith should be based on evidence rather than presumption or authority[6]. I am Christian because I firmly believe that some form of Christianity is true.

[6]Note that the commonly cited definition of faith as 'belief without evidence' is a modern invention only popularised in the nineteenth century. When a knowledgeable Christian discusses faith, they mean evidence-based belief and trust used to extrapolate towards some future (or otherwise as yet unknown) circumstance. It means gathering evidence, drawing up conclusions from that evidence, extrapolating to an unknown circumstance, and being prepared to act accordingly. For example, evidence may convince the Christian

'Some form' is, of course, a reflection that modern Christianity is diverse; my own Catholic wing of classical Anglican Evangelicalism just one of many streams. My writing is intended as non-denominational; I believe that what I say is consistent with all the creeds, confessions and catechisms of classical and orthodox Christianity, whether oriental Orthodox, Eastern Orthodox, Roman Catholic or Protestant (though not with some more modern expressions of and deviations from Christianity), although I will express things in the language of the Latin Churches (Roman Catholic and Protestant), as most of what I know was taught to me by the Latin theologians and philosophers, and their successors. Of course, I might be mistaken, and some denominational bias could have entered my writing; I can merely warn the reader in advance to watch out for it.

However, I should also emphasise that this book is not based on Christianity; it is a book about physics and philosophy. While Christians are theists; it is possible to accept theism without accepting Christianity, or indeed any of the religions which have claimed to have been revealed. My intention is to advocate for classical theism; and while the traditional expressions of Christianity are built on some form of classical theism (whether Platonic or Aristotelian) not all classical theists are Christian; and indeed not all Christians are classical theists. My positive case makes few references to specific Christian doctrines; I do not assume any Christian ideas, but take this work to be a prerequisite for Christian theology rather than a consequence of it. The exceptions are my brief discussion of the Trinity in section 13.4, and in my presentation of ethics in chapter 14 my selection of applications of the ethical theory is taken from those principles which the Christian philosophers have given a particular emphasis to. One may agree on a list of virtues, but disagree on which virtues are most important, and any selection is likely to reflect one's own biases. My historical discussion of the rise of modern science also interacts with Christianity, since science arose in a Christian culture. However, my choice of objections (and my responses to them) refer more frequently to Christianity, in the sense that the objection might question whether a particular notion is consistent with Christian theology, or attack a particular aspect of specifically Christian theology. I do not raise any objection specifically targeted against Islam (or some other form of theism), not because I don't consider such objections important (on the contrary, since I am not a Muslim it should be obvious that I have numerous objections to Islam which I find convincing), but because they lie outside what I am trying to achieve. My responses to those objections will also therefore contain specific references to aspects of Christian theology.

Christians believe that there are three legitimate sources which can be used as a basis for theology:

(1) Reason built on premises derived from observation of the natural world.
(2) Divine revelation (that is, information received directly from God through special acts or in Person).
(3) Reason built on premises derived from the conclusions of (1) and (2) together.

The Church is (even if it has frequently failed) the guardian of the revealed doctrine, and responsible for judging whether (3) is carried out correctly. The Christian will claim that (1) and (2), when correctly deduced and interpreted, are not in contradiction, but in harmony. Indeed, almost all of the theological consequences of (1) are contained in (2); a Christian can derive as much as they need to know about theology from (2) alone. So why should the Christian bother with natural reasoning? Firstly, it can tell us many things which are not contained within revelation, for example much of modern scientific knowledge (which is of significant interest in itself), and by extending moral reasoning to new circumstances and situations. Secondly, there are certain presumptions that are required before the acceptance of revelation, not least that revelation is possible, and how to recognise it and evaluate whether or not claimed revelation is true revelation. These cannot be established through

---

of the existence, trustworthiness and some particular promises of God; the faithful Christian will then live their life assuming that those promises will be kept at the appropriate time (Hebrews 11:13-16;39-40).

revelation without descending into circularity, but instead require a different source, which leads us back to (1). Thirdly, revelation is recorded in words, and words are often ambiguous. There are many possible ways of interpreting them. With only the divine source before us, we can, and should, list all the valid interpretations given the textual, linguistic and historical context. But we need something else to judge between those interpretations; and that is where (1), (3) and the (primarily historical) judgement of the Church come in.

My intention in this work is to focus solely on point (1); to take natural reasoning, and go not as far as I can with it, but far enough that it can fulfil the purposes outlined above and can be used to demonstrate that natural reasoning and Biblical revelation are consistent with each other, though I will not do either of these here. I do not (intentionally) base any of my conclusions on scripture, but only on empirical observation and natural reasoning.

My goal in this work is to investigate the extent to which Aristotelian metaphysics is consistent with quantum field theory, and to draw out the consequences of my findings for classical theism. I begin, in chapter 3, with a brief overview of metaphysics in general, and in chapter 4 with an overview of Aristotle's metaphysics in particular, outlining especially the motivation of his system as a means of solving the problem of change and constancy, and how this leads to the concepts of potentia, form and final cause or an inherent direction of matter. In chapter 5, I mention a few features of Aristotle's physics, since this was the starting point from which the modern scientific revolution diverged. This section provides the essential background to understanding my thesis. Greek philosophy is not well understood in today's society, even by most professional philosophers who don't specialise in it. It is very different in its fundamental assumptions and premises from the enlightenment philosophy that shapes our contemporary society; to someone used to modern ways of thinking it can seem strange and unworldly. Of course, to the one trained in classical philosophy modern thought is strange and unworldly. Because the differences between classical and modern thought go down to the basic and fundamental premises, which are often accepted unconsciously, it is easy to bring in assumptions which have no place in the old system when thinking about it. When evaluating a foreign philosophical system, it is important to judge it on its own terms, to see its own internal self-consistency and logic; only then can you appreciate it and evaluate it. My introduction is only brief, taking out only those parts of the classical tradition which I intend to interact with. There is no substitute for reading the original works, or contemporary introductions to classical thought (I cite a few examples at the appropriate place).

The next section of my work deals with modern philosophy, which is perhaps moving onto more familiar ground; even for those who haven't specifically studied philosophy have absorbed many of its assumptions from the surrounding culture. Again, I cannot discuss the philosophy in depth, but I will only focus on the interaction with physics, which is my chief interest. I am particularly interested in the rise of the mechanical philosophy. Most of modern philosophy is either built on this philosophy, or exists in reaction against it (while still sharing many of its key assumptions and definitions), or some combination of the two. The mechanical philosophy justifies itself from the success of Newtonian physics (though it itself is considerably older than Newton), and many of its adherents thus erroneously proclaim it as scientific. We should not, however, confuse the scientific method with the philosophy that gave birth to modern science. The scientific method, of course, relies on various philosophical assumptions concerning nature; these assumptions are, however, rather general and, while inconsistent with some subjectivist, idealist or dualist philosophies, can still fit in well with many very different world-views. Mechanism adds to these assumptions. But the issue of induction remains: although the mechanical philosophy is consistent with Newtonian physics, it is harder to say that it is the only world-view consistent with it, and harder still to say that it is therefore true. Indeed, given that the Newtonian system is false, an inconsistency with it might be regarded as advantageous. However, in the eighteenth,

nineteenth and early years of the twentieth century (when the major Enlightenment philosophers operated) Newton seemed unassailable. The mechanical system itself, of course, comes in many different forms, and evolved in time; most notably the early idea that forces were due to physical contact needed wholesale revision after Faraday and Maxwell highlighted the importance of extended fields as well as localised particles. Nonetheless, there remains a common thread throughout the mechanical philosophy, and I do not feel that it is too over-simplistic to treat it as a whole as my main adversary.

The mechanical philosophy arose in late medieval Europe, and I am particularly interested in how it emerged from scholasticism and what arguments were used by its supporters in its favour and against the older classical system. It is worth reviewing this discussion in the light of quantum physics; and again, although my presentation is forced to be briefer than the subject deserves, I will highlight many of the thoughts of fifteenth and sixteenth century academia as classical philosophy was abandoned. However, I need to begin with a discussion of the medieval roots of modern science, which is the topic of chapter 6, showing how a combination of the merger of geometry and some aspects of physics, the experimental method, and increasingly obvious flaws in Aristotle's physics led to its overthrow. What should replace Aristotle was less clear in this period, though important progress was made in numerous different areas. I also discuss the early changes in philosophy, leading up to Descartes, in this chapter. The medieval scholars almost reached the concepts often attributed to Galileo; they almost made the crucial breakthrough, and they made it easier for the few scholars of the early modern period schooled in the best medieval scholarship (such as Galileo) to go the rest of the way. Unfortunately their progress was halted by first the black death, and secondly the Renaissance, which, with its disdain for medieval scholarship, was more of a hindrance than an advance to physics and mathematics (though obviously a leap forward in art, architecture and engineering). Chapter 7 resumes the story of the emergence of modern physics with Galileo, jumps forward to Newton, and proceeds to give a brief overview of the most relevant (to my overall thesis) advances in physics and mathematics up to Einstein's work on relativity. This provides an essential background to the thesis I will develop in later chapters: special relativity and electromagnetism are crucial contributions towards quantum field theory, and one cannot understand quantum field theory without understanding the classical theories of special relativity and Maxwell's theory of electricity and magnetism. The mathematical formalism behind these theories continues almost unchanged into modern physics; the same symbols are used, but with a different understanding of what they represent. General relativity is, famously, more of a problem to merge into quantum physics. I will briefly discuss this problem in chapter 15, but far from conclusively; a discussion of classical general relativity is therefore essential to help the reader appreciate what I view as the main weakness of my thesis: until we have a quantum theory of gravity established, there must always remain an element of doubt over any interpretation of quantum physics.

I conclude my discussion of classical physics in chapter 8 with a discussion of the philosophy of classical physics, and in particular the issues that classical physics raises for classical philosophy, and, consequently, the classical theory of God. I do not think that these issues are devastating to classical theology, but they are still suggestive, and indeed did make suggestions to scholars of the nineteenth and twentieth century; the rise of deism and atheism on the one hand, and liberal theology on the other (as an attempt to find a new understanding of God in terms of the modern philosophy; I task I take to be hopeless but which nonetheless has been pursued with vigour). It is important to note that, at least at the popular level, the modern science/religion debate is still largely dominated by modern philosophers arguing it out with modern theologians; each having accepted at least to an extent many of the assumptions of the enlightenment. What I want to propose is, however, that classical theism was not killed by science, and this will change the debate. Modern theology and modern atheism can easily dialogue because they share a common language

and common set of assumptions. It is generally believed that atheism is winning that debate, and I personally would share that belief. However, classical theism is very different; it is built on a wholly distinct set of assumptions, and thus the arguments of modern atheists seem as risible to the classical theist as the arguments of classical theism do to the modern atheist: each side believes that the other starts in completely the wrong place, and rapidly descends into nonsense. The reason why dialogue is so difficult between these two traditions is that they are built on very different premises – often held unconsciously – and it is very difficult to debate where there is no common ground. The intellectual chasm between modern and classical theism is no less broad, and the warfare in many modern denominations is due ultimately to this chasm and that neither side means the same thing when they use the same words; each side appears incomprehensible to the other (or they try to understand their opponents – and get it completely wrong because they cannot think in terms of the opposing world-view).

With my discussion of classical physics complete, it is almost time for me to turn to the twentieth century revolution that overthrew how we think about the world, the birth of quantum physics. First, in chapter 9, I review an idea, which although it was first developed in the nineteenth century (and thus in terms of classical physics), is at the centre of much of modern physics, namely statistical physics. This is the physics of large numbers of particles; too many for us treat each one individually. However, we can describe the ensemble of particles in terms of a distribution, and from this distribution compute many of the macroscopic properties of the theory. Statistical mechanics is crucial in our modern understanding of gases, fluids and solids; it allows us to easily move from a microscopic theory to a macroscopic physics. However, it also contains a few other similarities with quantum physics. Firstly, its mathematical form and structure is remarkably similar to that of quantum field theory, the theory behind fundamental physics, so much so that ideas from the physics of solids have often passed into particle physics and vice versa, even though it might seem that their objects of study are very different. Secondly, statistical physics gives us our first taste of probability as a part of physical theory. The theoretical uncertainty in quantum physics comes from a very different source to that of statistical physics: in statistical physics, there is simply too much data for us to measure everything precisely, so one has to treat that uncertainty in a systematic way; while in quantum physics the uncertainty is inherent. But again, the use of probability in both cases means that there is a similarity in formalism; statistical physics prepares us for quantum physics. Thirdly, statistical physics offers our first example of how a theory with wide ranging and profound implications can arise from a handful of simple axioms. Pre-Newtonian physics was mainly collections of experimental results, with theories derived by induction to describe those results. Newton introduced a greater degree of formalism into the system (and subsequent developments, especially those of Einstein, increased this further), in that a great deal of phenomena could be derived from his equations of motion; but those equations of motion were chosen mainly because they worked. The Maxwell-Boltzmann distribution continues this trajectory, with a smaller number of empirically derived axioms. Logic is the movement from premise to conclusion. If we are to link physics with philosophy, we need to reduce physical theory to a (hopefully) small number of fundamental axioms, which we can then link with a philosophical system. The Maxwell-Boltzmann distribution can be derived from a handful of basic assumptions, the most important of which are locality – that events far away from a particle only influence it indirectly through intermediaries – and the conservation of energy. The conservation of energy (in classical physics) can be derived from the idea that there are laws describing the evolution of particles, which do not change in time and can be derived from a least action principle. Clearly, we are not yet at a position where we are ready to interact with philosophy – that there are time-independent regularities in nature is a prediction of various metaphysical systems, but where does the principle of least action come from? (Indeed, you may be asking what it is – I will explain at the appropriate

juncture.) But we start to see how physics and philosophy might interact: philosophy provides the axioms; physics enumerates the consequences of those axioms.

So, in chapter 10, we finally reach quantum physics. Here the mathematics gets rather more intense (although I also deploy mathematics in my descriptions of relativity and statistical physics), which is, I fear, unavoidable. My aim is to show how axioms drawn from a particular philosophical system can be used to derive modern physics; however modern physics is written in the language of mathematics, so to demonstrate that my conclusions are equivalent to modern physics, I have to express them mathematically. That is, I have to express the axioms in an abstract symbolic language, which I can then relate to the theory of sets, operators and manifolds. My description of quantum physics proceeds in three stages: first of all, I rewrite Newtonian physics in the language of quantum mechanics. This actually has some historical precedence: the first stage in the overthrow of Aristotelian physics was its re-expression in the language of geometry that would later be used by Galileo and Newton. My purpose is principally to familiarise the reader with the basic formalism of quantum mechanics, introducing the ideas of discrete states, operators acting on those states and probability amplitudes, without leaving the familiar shores of Newton's physics. The second stage is quantum mechanics. Quantum mechanics is not the fundamental theory of physics; in some respects it is even more wrong than classical physics, since the Newton/Einstein picture is an approximation of the true theory valid as the limit of one particular parameter is taken smoothly to zero, while quantum mechanics is not valid in that sense. However, historically quantum mechanics came first, and in terms of the formalism it is half-way between classical physics and quantum field theory, so it is useful to move towards the full theory in stages. The main weakness of quantum mechanics is its adherence to one of the assumptions of the mechanical world-view, namely that matter is indestructible and cannot be created or destroyed. However, when the formalism of quantum mechanics is merged with special relativity, we find that quantum particles are, in fact, created and destroyed, indeed very frequently. This means that we need another approach to the problem, and that approach is the mature form of quantum physics, namely quantum field theory. Quantum field theory is (or, at least, our current understanding strongly suggests that it is) the fundamental framework of the laws of physics. I go on to describe the standard model of particle physics, which is our current best description of the basic nature of physical reality. The standard model is not the correct theory – it is at best the zero gravitational field limit of the correct theory (at worst, there are some additional missing particles that need to be added to reach that theory) – but I am confident that it is close to the final answer. My main assumption in this work is that incorporating gravity into the theory is not going to drastically change its structure. Quantum field theory is mathematically very beautiful – at least until you try to calculate observables which you can compare against experiment, when it becomes rather messy rather quickly. But that messiness is as much a result of the approximations we use to make the calculations; the structure itself is wonderfully harmonious.

Up to this point, the end of chapter 10, I believe that my work ought not to be regarded as especially controversial. There is nothing particularly new in what I write. There will be a few simplifications in some of the reasoning to reach the conclusions; a few points where, if my goal was to write a physics textbook, or history of physics textbook, I would have to be a little more rigorous and detailed. I acknowledge that I have taken some shortcuts to keep the book more focused. Some of my notation is a little unconventional. But my conclusions would, I believe, be acceptable to most experts in the relevant fields of physics, history and philosophy, whether atheist or theist. Chapter 11 is where I start to tread on new ground.

While getting to quantum field theory required a lot of hard graft by a large number of immensely smart people, and not a few blind alleys, now we are there we can see that it again arises from a small number of fundamental axioms. It also provides, as conclusions, various

statements about the nature of matter. Both the premises and conclusions of field theory can therefore be linked to metaphysics, and I intend to look at both sides of the matter. I start in chapter 11 by examining the consequences of the conclusions of field theory, comparing what it states about matter against the two dominant historical metaphysical realist positions, the mechanical and the Aristotelian world-views. These, I believe, represent the only reasonable realist perspectives (and the alternatives to philosophical realism are strongly challenged by quantum physics), and I find that neither matches up perfectly with field theory. However, the Aristotelian position is far closer to what we need than the mechanical. Almost every aspect of Aristotelian philosophy has some analogue in quantum field theory; rarely identical but close enough that we can match them. There are a few aspects of Aristotle's philosophy which need to be changed at a more fundamental level - most importantly, we need to merge into it the mechanical emphasis that space and time can be represented geometrically – but my claim in this chapter is that by merging together various aspects of the two world views we can arrive at a metaphysics in which quantum field theory makes sense. Thus we find that our philosophy can once again start to inform our physics. The interaction between the two goes the other way as well; one of the criticisms of the early modern philosophers was that key Aristotelian concepts such as form were not precisely enough defined. But, of course, the Greeks and medievals couldn't provide those definitions, but only work with the concept as best they could, because they didn't have access to quantum field theory to inform them of the microscopic nature of reality.

With the metaphysics in place, I turn my attention to its consequences, firstly considering the case of classical theism and what it says about the nature and existence of God. I begin in chapter 12, where I discuss a few historically important objections to the arguments used in classical theism, and discover that they don't really do what they are often claimed to do. Classical metaphysics is different from modern metaphysics, and as a consequence the standard arguments for the existence of God differ subtly as well. The modern objections to the classical arguments use modern philosophy to try to respond to arguments for God, with the premises of modern philosophy. This means that they might accurately dispose of arguments for God posed by modern philosophers, but if deployed against classical arguments the contradiction in the premises used by those objecting to the arguments with those who made the arguments mean that there is a mismatch: any difficulties found might well arise from those contradictory premises. And thus we find that the standard arguments against the classical cosmological and teleological arguments fail. A straw man was erected in place of the cosmological argument. While strong arguments were raised against the argument from design, this differs from the teleological argument in important ways, and thus those objections either have no force against the teleological argument or tend to support it. Kant was even worse: he offered a different definition of God to that employed in classical theism, so his objection that the cosmological argument didn't reach his definition of God is irrelevant: it wasn't intended to. The best modern objection against the classical arguments that I am aware of was by Kenny (Kenny, 1969) (who was more careful than most to compare the various versions of the argument and present them in their best form), but even this falls short: he argued that the arguments are fatally dependent on an outdated cosmology. Inasmuch as he is referring to Aristotelian physics, this is false, because although the classical writers used Aristotelian physics to illustrate their arguments, the arguments themselves do not depend on this physics. Inasmuch as he refers to Aristotelian metaphysics, this is false, because the metaphysics is not outdated.

In chapter 13, I then build on the metaphysical principles shown to be consistent with modern physics, to develop a classical theory of God. Beginning with the definition of God as a terminus in a series of efficient causality, I demonstrate God's existence and go on to show how the standard attributes of God such as his simplicity, eternity, intellect and will follow from the underlying metaphysics, physics and definitions. I also answer various common objections to these attributes or the existence of God, and show that they are either based on

misunderstanding the arguments, introducing incorrect modern metaphysical assumptions and premises, or fallacious.

The second consequence of the metaphysics I wish to discuss is ethics, outlined in chapter 14. I concentrate on two objections to ethical theory, Hume's is/ought dilemma and the claim that to define goodness is to commit the naturalistic fallacy by Moore, and show that both these objections rest on various premises, most importantly the rejection of the Aristotelian concepts of formal and final causality. The consequence of this is that 1) Any ethical theory built from the rejection of formal and final causes, which includes consequentialism, human rights based approaches to ethics, and every other modern approach to ethics except for Kant's (and those related to Kant's approach) is incoherent; 2) Kant's system also fails because his categorical imperative drawn from the requirements of logical consistency and universal applicability is not sufficient to determine an unique ethical system; 3) Aristotelian approaches to ethics avoid the implications of Hume's and Moore's arguments because they reject their premises and instead insist, and as I show in the previous chapters rightly insist, that form and finality are essential parts of the definition of any physical object. Final causality leads to a natural definition of goodness as a target which objects ought to aim towards becoming, with other targets being contradictory to the nature of the being, and the link from final causality to formal causality provides the link from 'ought' to 'is' that Hume failed to consider (note the order: Hume tried to argue from 'Is' to 'Ought' and failed; the Aristotelians take 'Ought' to be the more fundamental principle, and argue from that to 'Is'; a direction forbidden by Hume's empiricism, but in other systems, including the scientific approach, perfectly possible). From this, I briefly outline some of the main features of the developed form of Aristotelian ethics, called *natural law ethics*. This approach is sadly neglected in our contemporary age, but remains the only known candidate for a rational objective ethical theory.

Thus far, I have constructed a metaphysics that is consistent with modern physics, shown that an alternative metaphysics which historically replaced it is inconsistent with modern physics, and drawn conclusions from it for theology and ethics. But, while physics is prior to metaphysics in the order of our knowing, in the natural order of knowledge metaphysics is more fundamental. So can we draw conclusions about the nature of physics from the conclusions of the previous chapters. If we can assume the existence of God, what does that tell us about the nature of physics? Starting in chapter 15, commencing with premises drawn from classical theism, I begin drawing out a line of reasoning about how God interacts with the world that will lead to modern physics as a description of God's acts to sustain the world under the approximation that God is impartial to mankind (and other beings in the universe), in the sense that His actions are not affected by any good or evil in the world. God's freedom to act implies that that physics ought to be probabilistic rather than deterministic; God's timelessness and universality imply that the theory is nonetheless constrained by various symmetry requirements. From these basic principles, I can build up a general framework, with our current understanding of particle physics one particular member of the set of possible physical theories. If we constrain the theory further, by supposing that there need to be stars around for long enough for evolution by natural selection to occur, that there is complex chemistry and the basic requirements for the existence of complex life forms are satisfied, then the range of possible physical theories is reduced considerably. Only a limited number of the possible physical theories are consistent with the existence of rational animals such as ourselves able to form relationships with God. Thus if God intended to create such beings, His options were limited. This gives theism far greater predictive power than naturalistic metaphysics; and given that the universe we observe is in the range consistent with theism, this makes theism a far more likely explanation of current physics than some anti-theist philosophy (even if we leave aside the arguments of previous chapters suggesting that the most natural metaphysics consistent with physics implies theism). We cannot (or at least I cannot) say that assuming the classical picture

of God, physics has to be the way it is, but I can come pretty close. I also discuss, more speculatively, how quantum theories of gravity might fit into this picture.

Of course, if theism is true (or at least the truth resides somewhere among the most common expressions of theism), the idea that God is impartial to the fate of mankind can be approximate at best. If we relax this approximation, then that would imply a localised breaking of the various symmetry and locality conditions that were a conclusion of the previous chapter; since these conditions lead to the conservation of energy and momentum, this means that we might expect to see events in the universe where the usual conservation laws don't apply. This is not contrary to our understanding of physics, but a consequence of the more fundamental principles that led us to physics. I argue in chapter 16 that this gives us a way to define and identify miraculous events; and to demand, given the conclusions of the chapter on ethics, that miracles ought to have occurred in the history of mankind. Of course, in a naturalistic framework (as most famously argued by Hume, but also other writers, each bringing in assumptions and definitions into their arguments that the theist would reject) the miraculous is impossible. Thus if miracles occur, then atheism is false. The difficulty is in judging whether or not miracles occur given that our interpretation of the evidence depends on various assumptions which we bring in from our underlying metaphysics, and thus the previously determined theist would look upon that evidence more favourably than the one who had previously accepted atheism. Evaluating the evidence to assess whether a miracle happened from a neutral point of view is impossible, because we need to bring in assumptions to interpret historical data and ultimately the set of assumptions agreed upon by both atheism and theism is not sufficiently large to make that assessment. I therefore propose that instead of interpreting from the evidence to the miracle, and then from the miracle to atheism or theism, we instead ask what is the probability of that raw evidence arising conditional upon either theism or atheism being true. This allows us to use the evidence for miracles (or lack of miracles) as a means of judging between theism and atheism. Such evidence would show that whatever is behind physics is not impartial to mankind, and would therefore act as evidence against a key presumption of atheist metaphysics. I also suggest that such evidence is far more prevalent and far stronger than many in the modern world realise.

Of course, I end in the place where many people would begin. I do not in this work present arguments to allow us to choose which religion (or irreligion) is true; but that is fine, because it is not my intention to go that far in this work. What instead I have done (or at least attempted to do) is to construct a framework, with premises consistent with fundamental physics and conclusions constructed from built upon sound philosophical and rational principles, that allows us to at least address the question. In the process, I hope that I will have demolished a few of the myths and misunderstandings of our modern age. Of course, I cannot show that my theistic model is the only model consistent with physics; although I can show that the line of thought that led to modern atheism is erroneous. Is there a naturalistic metaphysical picture, an alternative to the mechanical world view, which can do as well as the theism I will propose here? That, I think, is for the atheists to put forward. Nothing that I have seen thus far is up to the task.

Very little of the work presented here is original to me – only in chapters 10 and 15 have I consciously deviated from the research of others and introduced my own material - although hopefully I present even the older material in a fresh way. My main contribution is to tie together what have until now been regarded as diverse fields of study. I do not claim that that my work here is the final word on the subject; not even my own final word. In a book of this scope and length, it is inevitable that there will be objections I haven't considered, or places where my argument needs clarification, more fleshing out, or correcting. However, I believe that even if some of the details of what I have written need polishing, the general thrust of my argument will, in time, prove fruitful and productive. I hope that this work will be of benefit to others besides myself.

This book is intended for the intelligent but non-expert reader; indeed it covers a lot of ground over many different subject areas. I am therefore not able to go into any topic in the depth that would be ideal. I have tried to introduce every concept so that a reader who is not familiar with the subject (and I think that almost all readers will be non-expert on at least some of the fields I write about) can nonetheless understand the thrust of the argument and where it fits in with my overall thesis. However, some of my description of physics, particularly my discussions in 10 and 15, where I have no choice but to delve into the mathematical language in which physics is most properly expressed, requires knowledge of mathematics up to advanced level school studies (for example basic calculus, matrices, and so on). While I briefly describe what the various mathematical symbols represent, my introduction is unlikely to be sufficient except as a reminder for those who have some prior study of mathematics. My reason for including this mathematics is because I am incapable of describing modern physics accurately and concisely without using the mathematical structure it is best expressed in; and for this project accuracy is important. To keep the work focused, I have skipped over a lot of the details in the descriptions of quantum field theory and classical philosophies. I have provided the minimum required to make my point and (hopefully) be understandable. That does not mean that the details I have omitted cannot be provided; only that to do so would be a distraction from my main purposes. They can be found in the standard textbooks of advanced physics or classical philosophy. I therefore do not expect this work to be an easy read. I do, however, hope that it will be a rewarding one.

In any work of this length, some mistakes are inevitable, especially in the minor details in subjects which depart from my own specialisation of physics, and those areas of the philosophy which I know best. Here I might (through my ignorance) present contested details as though they were fact. For these errors, I can only apologise. Obviously I have done my best to avoid mistakes, and to represent the views of my opponents fairly. I do not believe that I err in any major fact which would damage my thesis; but if I am mistaken anywhere, I beg the readers indulgence and request that they inform me so that I can make the appropriate corrections.

CHAPTER 2

Tools

2.1. Logic

Logic is a means of arriving at conclusions which necessarily follow from at least two premises. There are three steps in a logical argument: 1) the collection of premises; 2) an argument combining those premises; 3) the drawing of a conclusion. The claim is that if the premises are correct, and if the argument is correct, then the conclusion must also be correct. If the premises are incorrect, or the argument incorrect, then the conclusion may be either true or false; though it will usually be false. It is possible to obtain a true conclusion from incorrect premises or an incorrect method (for example: premise (a) $1+1=3$; premise (b) $3+2=4$; method: substitute in the expression for 3 from premise (a) in place of the 3 in premise (b); conclusion $1+1+2=4$: the conclusion is correct even though both premises are nonsense), though this doesn't happen very often. Thus casting doubt on a logical argument does not in itself prove that the conclusions are false: one also needs a valid chain of reasoning leading to a contradictory conclusion. It is therefore essential to state the premises alongside the conclusion after any rational analysis, otherwise those hearing the claim cannot hope to judge what it means.

If the conclusion of a logical argument is found to be wrong, then either the argument must be wrong or at least one of the premises must be wrong. In practice, invalid logical arguments are presented far too frequently, though in principle, with enough care, they can be avoided or at the very least noticed. The argument proceeds according to certain well defined rules, and all we have to do is carefully list each rule we use at each step of the process. Everybody makes mistakes, but usually people make different mistakes (or the same person makes different mistakes when trying to reproduce his argument), and thus through a careful process of checking, and enough time, errors in argumentation are almost invariably found and corrected. Assuming that this process of checking is done. That is not always the case. Argumentation should therefore be presented in a way that makes the process and premises clear, to make life easier for those who need to review and check it.

The premises are, however, more easily disputed, and it is harder to prove them to be right. Ultimately, the most we can say through logic is not 'C is a true statement', but rather 'C is necessarily implied by the premises P_1, P_2, and so on.' A separate argument, or observation, is required to establish each of these premises. On the other hand, if the conclusion is right, that does not necessarily mean that the premises are correct. Sometimes different and in some respects contradictory sets of premises can lead to the same conclusion.

There is one way in which we can be certain of the premises given a true conclusion, though it is only applicable in some circumstances. Suppose that there are two sets of propositions, P and C. We take P to be premises, and C to be conclusions, and correctly reason that P logically implies C. We then confirm that a subset of C, $\bar{C}$, is true by comparing with experiment (this is possible if C refers to various necessarily discrete qualities); the remaining members of C, $\bar{C}'$, are either untested, or not accessible to experiment, or the experimental results are ambiguous; they are not proved false. As discussed above, this by itself gives us no firm indication about whether the propositions P are true or false. However, if we can then establish a valid chain of deductive reasoning that takes $\bar{C}$ as premises, and concludes that if $\bar{C}$ is true, then P is true, then we have established the truth of the

premises P, and the falsity of every other set of premises that lead to the same conclusions. We have also demonstrated the truth of $\bar{C}'$. We need not even go as far as this; if we can reason from $\bar{C}$ to $\bar{P}$, a subset of our original set of premises, that might be good enough if the remaining premises $\bar{P}'$ are self-evident, such as the principle of non-contradiction, or can themselves be tested experimentally. It is thus best, if possible, to construct arguments in both directions.

That the conclusion follows from its premises by a correct chain of reasoning does not mean that it is true. All truth is reasonable, but not everything reasonable is true. Even if the conclusion is reasonably derived from its premises, and agrees with all known observations, that does not guarantee that the premises are correct, since there might be a contradictory set of premises which do equally well. Proving something false is easy: one simply has to show that either the conclusion or one of the premises contradict observation. Proving something true is immensely harder. However one can reduce uncertainty of the truth to within some narrow bounds, or perhaps a few discrete sets of narrow bounds, with everything else eliminated.

The conclusion draws out information that is not visible from any one of the premises alone, but only from them in combination. If the conclusion can be shown to be identical to one of the premises, then we do not have a logical argument, but a tautology. A tautology is true, and may be useful in some senses (for example, it may be useful to rephrase a statement in a different but equivalent way which makes some of its implications more obvious), but does not give us new information. A true logical argument, however, increases our knowledge because it draws out more information than is contained in any of the premises alone.

Premises may arise from various sources, which include:

(1) Self-evident propositions;
(2) Definitions;
(3) Propositions which arise as the conclusions of other arguments;
(4) Propositions which are derived from observation and empirical tests;
(5) Hypothesis and guesses which we desire to test by examining the conclusions which we draw from them;
(6) Assumptions (often unstated) that follow from a particular world-view.

The most important of the self-evident propositions is the principle of non-contradiction: that something cannot both be and not be in the same respect at the same time, or that a statement cannot be simultaneously true and false. This principle cannot be proved by logic; rather it is the basis of logic.

Sound definitions are also key to any logical argument. A definition must be clear. It should ideally distinguish firstly what something is, and also what it has that everything else lacks. In particular, one common problem occurs when the same word is used to convey different meanings. In this case, one should always be careful to state which of the meanings is intended at each step of the argument. Definitions need not, and for most physical objects should not, be perfectly precise. There could well be some allowed variation. This is particularly the case when we are making a definition of a group of individual beings, or we are considering a being which is capable of change. For example, when we define the word 'tiger', we do not visualise one single tiger, and leave it at that, because that particular single static image will not match anything in reality. Rather our definition should be looser, listing what aspects of tigerness are held by all tigers, and how tigers should be distinguished from everything else; or perhaps a rule to compute every possible state of matter which corresponds to a tiger while excluding all the non-tigers. Our definition should encompass tigers from the moment of conception, when the new life is formed, to the tiger in extreme old age. A definition is not simply a list of properties (such as colours, sizes, and so on) which tigers have; as though the essence of something were merely a bundle of properties. After all, a single tiger will exhibit different properties at different stages of its life. Rather, it should represent some principle from which the properties can be derived.

I will define a logical system as a consistent set of definitions, self-evident propositions and other axioms, and the full range of conclusions that can be drawn from these.

One of the most important results of the twentieth century was a theorem deduced by the Austrian Mathematician Kurt Gödel. This says that no logical system is complete. In other words, within every logical system there must be at least one axiom which is not provable within the system; and the self-consistency of the system cannot be proven from within the system. This means that, contrary to the hopes of certain philosophers, nothing can be proved by logic or reason alone. We must introduce other evidence, usually from our senses. If we believe that reality, that is the description of all that is observable (either in practice or only in principle) or otherwise intelligible, is complete (i.e. requires nothing outside itself to be intelligible), then we have to conclude that reality is not a system of logic. That does not mean that reality is inconsistent with logic, or that logic is of no use: on the contrary it is of the utmost importance. We can proceed by capturing certain aspects of reality, formalizing them either mathematically or in terms of other logical statements, and drawing conclusions about certain other aspects of reality. This process allows us, for example, to make predictions about events which we have not observed, perhaps because those events are in the future, or to understand the links between various events and how one event causes another, or how one being is the cause of another being. If done correctly, the formalism we use to do this captures some aspects of reality. Perhaps a great deal of reality, or even most of it. But we must never confuse this formalism with reality; it can never capture all of it. For each formalism is a logical system, and each logical system, unlike reality, cannot be complete.

Logic is thus exceptionally useful, but is not everything. There are two errors then to be avoided. The first is to think that any logical system (whether it is some philosophical or scientific theory) can explain everything. It cannot. There is always something beyond the reach of science. This error, which is manifestly inconsistent (since the statement 'The only statements we can know with certainty are those which can be validated through science' is itself a statement which is treated as though it were certain, but one which cannot be validated through science), is, unfortunately, far too common in our present age. The second, and far worse, mistake is the opposite extreme, and to think that nothing can be studied logically; that we should rejoice in inconsistencies and paradoxes. We should never use the precept that because logic is incomplete, it can be ignored. Everything ought to be studied using logic as much as it can be; it is among the most powerful tools we have.

There are three basic types of argument available to the logician: the deductive, the inductive, and the comparison. The deductive argument is when one attempts to move from premises to conclusions. These are to be preferred, since, if the premises are correct, there is no room for any mistake. The inductive argument attempts to uncover the premises that might give rise to a particular conclusion. This is more dangerous, because often any solution found is not unique. One may successfully uncover one set of premises, but no matter how hard you work, it is often very difficult to be sure that you have found every possible set of premises that are consistent with that conclusion. The third method is comparison: one finds one conclusion X and another conclusion Y, and can ask whether they are the same as each other, or perhaps quantify how different they are from each other.

More importantly, comparison is used to compare a logical system with observation. One can, for example, make predictions (i.e. a conclusion) from a particular kind of logical system, a theoretical model, and compare them against observation. Or one can sometimes take an observation and compare it against the premises of a theoretical model. What one cannot do is directly compare the observation of one event with another observation of another event. For example, suppose that I observe the sun rising today, and also the sun rising yesterday. How can these events be compared, except through some model? The observations themselves are just two distinct blips on my retina. To make a comparison, I need to extract certain properties observed in each event; and I can say that there are

certain similarities and differences in those observable properties between the two events, but as soon as I extract and label properties, I am moving away from observation and into formalisation and therefore modelling. Even if I extracted every possible observable property, that collection of properties would not be the observation itself but a description of various aspects of that observation. Even then, what of these properties? 'The sun today was yellow, and the sun yesterday was yellow.' So what? What connection is there in those two statements – unless I wish to make a model such as 'The sun was the same colour on those two days' or 'The sun was a different colour on those two days.' These statements are not observations: they are propositions which can be either be confirmed or rejected by comparisons with the observations.

Therefore theoretical models are of the utmost importance if we are to try to make connections between different observations. Another common error (for example, made by the empiricists) is to deny the importance or knowability of such models; to think that observations alone are enough to allow us to increase understanding. But now we run into a different problem: which model do we use? There may be many models consistent with the data. When drawing a conclusion from the observed data, we must be careful to understand how much of our conclusion is related to the observation, and how much is derived from the particular model we use (perhaps unconsciously) to measure and interpret the data. Almost every empirical conclusion we make is in some respect model-dependent. Indeed, the philosophy of the extreme empiricists[1] (who hold that only what is observed is knowable) is self refuting, because its premise 'The only facts we can know are observed data' is postulated as a fact but is not in itself observed. This does not mean that these conclusions are wholly unreliable, but it does mean that they contain a certain degree of imprecision which we should always try to quantify and state as well as we can. This imprecision is known as a modelling systematic error.

It is also wrong to put too much confidence in theoretical models. A model is almost certainly not a true representation of reality, but only at best an approximation to it. It will be built on a series of assumptions and often simplifications. It is therefore important to test models against new observations which were unavailable to those who formulated the model, and not begin to trust any model until they regularly pass such tests without having to be modified or altered.

Thus the best approach is one that combines both theoretical models founded on logic (or related subjects such as mathematics) and data gleaned from observation. Neither extreme is a sound methodology.

One argument that is sometimes made against the use of logic when comparing against reality is the claim that we cannot draw any conclusions that are not contained within our initial premises. One of the most basic tools of logic is the syllogism. A common example of the syllogism is

(1) All men are mortal.
(2) Plato is a man.
(3) Therefore, Plato is mortal.

We have two premises: one concerning all men, and the other concerning Plato, from which we draw the conclusion. It should be obvious that the conclusion is true necessarily if the premises are true.

The objection I want to discuss here is what I shall call the first empiricist fallacy. This is to suppose that the only way that we can see whether all men are mortal is to go and investigate each man in turn, and ask whether or not they have died. If they all have died, then we can conclude that all men are, in fact, mortal. But as part of this process, we

[1]Note that one does not have to be an empiricist to value the importance of empirical data. Empiricism is the name given to one particular philosophy; but it is not the only philosophy that values the empirical (or information derived from observation), but rather empiricism is the philosophy that denies anything that is not empirical.

also studied whether or not Plato had died when his turn came up. In other words, the conclusion is contained in the first premise: it tells us nothing new, and the second premise contributes nothing to it. Thus, according to the people who adopt this objection, this is a mere tautology rather than a logical argument.

The flaw in this objection is the assumption that we can only deduce that all men are mortal through observation of all men. Indeed, we cannot deduce that all men are mortal through observation. We cannot, for example, say that all men have died, because those alive today have not yet died, and certainly we can say nothing about future generations. We should instead note that a step was missing in the syllogism: the definition of what is a man. If the definition is clear and accurate enough, then we can in principle use it to deduce certain properties that all men satisfy. The definition will tell us everything that all men have in common, and also everything that men have that things which are not men lack. We do not need to observe any man to derive these properties from the definition (in principle, although in practice we are better served by testing our conclusions through observation of a few samples). One of the properties which can be derived from the definition is mortality. Thus we can say that 'All men are mortal' without requiring to observe all men; indeed without needing (but only in the strictest, almost impossibly challenging, sense of the word *need*) to observe any men. The second premise comes through comparison; we compare what we can observe of Plato against the definition. We find that Plato does not lack anything common to mankind, and possesses those things which everything that isn't man lacks. We can thus conclude that Plato is mortal without having seen whether or not he has died.

How do we know that we have defined 'man' correctly without observing any man? We don't: just because we define something does not mean that that definition is useful; that is, applicable to something in reality. But that ultimately is not of importance: we can continue making definitions until we find one that is applicable to the circumstance we want to apply it to. Obviously, in practice, we take a short cut, and use a few observations to limit the scope; then, with a good theoretical understanding, we can construct a definition that is useful for this syllogism. It is not necessary to observe all men to construct the definition of man; we need only study a few men.

Obviously, in biology it is difficult to draw the sort of conclusions we require from a basic definition, so I will offer another syllogism, where the process is hopefully more obvious. If the reader wishes to be more accurate and pedantic, they may preface each point in the argument with 'In Euclidean geometry'. I also assume the standard mathematical definitions of addition, equality, division, multiplication and squaring.

(1) A triangle is defined as a shape consisting of three points connected by straight edges.
(2) For any triangle, we may label the length of each side as a, b and c, and the angles opposite the side as A, B and C.
(3) Suppose that we have a triangle where B is a right angle. The cosine of the angle A is then defined as the ratio of the lengths of the two sides adjacent to the angle, $\cos A = c/b$. This is a mathematical function of the angle, and can be extended to all angles, whether or not they are related to right-angled triangles.
(4) All triangles (not just right-angled triangles) satisfy the relation $a^2 + b^2 = c^2 + 2ab\cos C$.
(5) $\triangle$ is a triangle.
(6) Therefore $\triangle$ satisfies the relation $a^2 + b^2 = c^2 + 2ab\cos C$.

The first three points are definitions of the triangle, notation, and cosine. The last three points are the syllogism. Now, it is important to note that we do not need to study every possible triangle to derive the cosine rule. Indeed, we do not need to draw or even visualise any triangle to prove it. Nor do we need to measure the angles and lengths to conclude that something is a triangle. In fact, just by observing that it consists of three points joined by straight edges, we can deduce that it satisfies the cosine rule without making any further

observations. In this way, the syllogism gives us new information not contained in either one of the premises alone, and is not a tautology.

In other words, there is no reason on this basis to conclude that we cannot use logic to draw at least some conclusions about reality which have not yet been observed.

I should also make a distinction between traditional and modern forms of logic. Modern logic (as developed from the end of the nineteenth century by the likes of Wittgenstein, Russell, and Whitehead) tries to systematise the whole process using tables and a pseudo-mathematical notation[2]. Traditionally, while mathematics and logic require similar skills and ways of thinking, they are seen as separate, in that they are applied to different situations. Traditional logic is seen as exploring the consequences of causal links, while mathematics is seen as being related to number, length, displacement and so on, i.e. discusses non-causal relationships. For example, consider the statement that 'Ducks are birds with webbed feet.' The traditional logician would consider this as a (partial) description of the formal cause of a duck. The modern logician would break this statement (labelled $B\&W$) into its two parts, namely that 'Ducks are Birds' (labelled proposition B) and 'Ducks have webbed feet' (labelled proposition W), and draw up a truth table:

B	W	$B\&W$
T	T	T
T	F	F
F	T	F
F	F	F

where T means that the proposition is true, and F that the proposition is false. What this table tells us is that the statement 'Ducks are birds with webbed feet' is true if and only if the propositions that ducks are birds and ducks have webbed feet are true. Which is all very good, but doesn't really help us understand our poor duck. Thus by making the whole subject more abstract, the logician has moved it away from reality. This statement was intended as a guide to help us define or identify ducks; the modern logician treats it as an abstract problem, a question of truth and falsity, completely isolated from the existence of any ducks on our nearest pond. The modern logician assumes knowledge of the duck prior to knowledge of its form. On the other hand, the traditional logician considers the idea that the proposition that 'Ducks are birds' could be false as absurd, since ducks are, by definition, a type of bird. The statement 'The proposition that "Ducks are birds" is false' is not so much a falsity, but more importantly a nonsense self-contradiction. Everything baring the first line of the truth table is thus not useful a guide to our reasoning, but utter nonsense. By reducing this statement into abstract form, the modern logician has missed the whole point.

---

[2]Note that I am being slightly misleading here in my terminology. There were three main periods (at least within Europe) when logic was studied extensively. The first was ancient Greece, starting with the early pioneers such as Zeno, and moving through the likes of Aristotle, until it diminished in importance as the Roman era began and there was little new development of any note. The second was the high medieval period, where the scholastics concentrated on nothing but logic; this was diminished in the renaissance where an over-emphasis on empiricism took the place of the earlier over-emphasis on reason. The third was in modern times, starting from the nineteenth century, where the success of mathematics encouraged philosophers to try to treat their subject in a more formal way. There were two forms of logic developed by the ancient Greeks. One of these was by Aristotle, and became the basis of what I am calling traditional logic. The other was developed by the stoics, most particularly Chrysippus, and was far closer to what I am calling modern logic, being concerned with the truth and falsity of propositions and combinations of propositions. However, during the Roman and most importantly medieval periods this stoic logic was largely abandoned in favour of Aristotle's system. Thus by traditional logic, I largely mean the system that was accepted in the medieval period. Modern logic was developed independently of the older Greek insights during the late nineteenth century as an attempt to bring the power of mathematics into philosophy; however it later was shown that there are clear similarities between it and the stoic logic.

Of course, this example is specifically constructed to illustrate the flaws of modern logic. In reality, no logician would think in this way. But sometimes the problems are more subtle. This sort of error of trying to apply modern logic to situations which are more suited to traditional logic can happen.

The difference is even more pronounced when we consider conditional statements, such as 'If I run, then I will get to my meeting on time.' The modern logician will split this up into two statements, namely whether or not I am running (proposition R), and whether or not I am getting to my meeting on time (proposition M), and will draw up the following table:

R	M	$M \setminus R$
T	T	T
T	F	F
F	T	T/F
F	F	T/F

where $M \setminus R$ represents the conditional statement, M given R[3]. Strictly speaking, the conditional is undetermined if R is false, so I have written the final line as T/F to indicate this uncertainty. However, this wholly misses the point. In traditional logic, the statement is an expression saying that my running is the cause of my getting to the meeting on time. Modern logic completely ignores this causal link; its main weakness is that by reducing the arguments to symbols we miss out on what sort of causal relationship (if there is a causal relationship) lies between the various aspects of the argument. It is only a little simplistic to say that modern logic is concerned with the relationship between entirely internal concepts (thus in principle the concepts 'duck' and 'bird' can be separated from each other), while traditional logic is concerned with objective relations between facts inherent in the external world (so the facts of the 'duck' and the 'bird' can't be separated). Traditional logic assumes material and metaphysical realism; modern logic doesn't care about the first and assumes the second is false. Traditional logic is limited in conception but broad in real world usefulness; modern logic broad in conception (since the symbols are not tied to meaning) but of limited usefulness when applied to objective reality.

To say, then, that modern logic represents all logic presupposes a Humean anti-causality which we should not assume (mainly because Hume's writings on causality were highly questionable, as I shall discuss later). That doesn't mean that modern logic is not useful in the right circumstances; it is an immensely powerful tool for those problems to which it is suited. But it does mean that traditional logic (based around the syllogism and causal relations) should be seen as distinct from modern logic and taught alongside it (rather than, as happens too much today, modern logic is taught instead of traditional logic), and each used in their appropriate circumstances. In this work, traditional logic is more appropriate, and therefore I will focus on it. Modern logic deals with relationships between abstract objects. The abstract has to be linked to the real through some sort of mapping. Traditional logic deals with relationships between real objects, and can use real and observed objects directly as its premises and conclusions. It is interested not only in possible causes, but actual causal relationships. I don't use a symbolic logic because the symbols will distract from what I am trying to say.

2.2. Probability

Another subject about which we need clarity is the study of probability. What is probability? At one level, the answer to this question is easy. We consider some indeterminate experiment, and enumerate all the possible outcomes of this event. To each outcome, we

[3]Conditional statements in probability or logic are usually denoted by an unbroken vertical line; but were I to use this symbol it would clash with some of my other notation, so I will use the \ symbol to denote a conditional statement.

assign a number between 0 and 1, ensuring that the sum of these numbers is fixed as 1. I will label this set of numbers as the probability distribution. Were we to perform the experiment an infinite number of times, then if the probability distribution is correct, the ratios of the total number of times each outcome occurred would be the same as the ratios of the associated members of the probability distribution.

This definition is obviously problematic in practice: we cannot perform any experiment an infinite number of times. There are indeed some experiments which we can only perform once, since we can never precisely reproduce the conditions under which it was first done. Nonetheless, we can still assign probability distributions for each experiment. The probability distribution is thus a useful theoretical construct or model which we can use to parametrise uncertainty. It can be studied and understood using various mathematical tools. In particular, the mathematical theory of probability is based on four premises:

(0) For every observation to which we can apply probability theory, it is possible to express the outcomes of that observation in terms of a set of irreducible, complete and orthogonal states. By irreducible, I mean that the outcomes cannot be expressed in terms of simpler parts; by complete I mean that it is not possible to express an answer to this question which is outside this set of states; and by orthogonal I mean that the answer cannot be in two different irreducible states simultaneously. The set of states might be infinite (for example, the outcomes could be something which can be mapped onto the set of all real numbers). For each of these states we can assign a real number called the probability.

(1) The probability for each individual outcome is greater than zero (the positivity condition).

(2) The sum of all the probabilities across the complete set of outcomes is one (the normalisation condition).

(3) The probability of the answer being in some subset of outcomes A or some subset of outcomes B is, if A and B are orthogonal, is the probability of A plus the probability of B (in symbols[4] if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$; the additivity condition).

I have labelled the first premise as premise 0 firstly because it is often left unstated in presentations of probability, because the other three premises rely on it, and to make it more memorable because it will be of particular importance later in my argument[5].

There are five interpretations of probability commonly encountered:

(1) **The classical interpretation:** Here we count each possibility, and assume that they are evenly distributed. The probability distribution is thus the number of ways in which we can construct a particular outcome. For example, if we take a balanced six-sided dice, with each side labelled from 1 to 6. When we roll the dice, we observe the number that comes up. Each outcome is equally likely, so we assign to each a probability of 1/6. Now if we had two dice, and were interested in the sum of the two results from each die, then we note that the possible results are each number from 2 to 12. But these are not equally likely. There is only one way we can achieve the result 2 (each dice shows the number 1, while there are two ways in which we can achieve the result 3 (dice A shows 1 while dice B shows 2, or dice B shows 1 while dice A shows 2). Thus we assign to the result 3 a probability twice that of the result 2, 1/18 compared to 1/36 (as they work out to be in

[4] $\emptyset$ represents the empty set, which has no members. $\cap$ represents the intersection operator, thus $A \cap B$ represents those elements which A and B have in common. $A \cap B = \emptyset$ states that A and B have no members in common. $\cup$ is the union operator, so $A \cup B$ represents the set of all the members of A and all the members of B.

[5] In practice, there are several different formulations of the fundamental axioms of probability theory, all ultimately equivalent. I use these definitions, called Kolmogorov's axioms, because they will be the most convenient starting point for my later work.

practice). All we do to obtain the probabilities is to count up how many ways in which a result could be achieved. If we wish, we can then perform the experiment by rolling the dice an infinite number of times, and we will find our result verified, if the dice are indeed perfectly balanced. We can extend this calculation to three, four, or however many we like dice rolls. Thus the classical interpretation is that probability is the enumeration of the number of ways in which an outcome can be achieved. We do not need to roll the dice to compute the expected probability distribution.

As a theoretical physicist, I have considerable sympathy for the classical interpretation, because ultimately it comes down to symmetries. A symmetry is a transformation of the system under investigation which leaves it in an identical state with regards to everything of significance (in this case, with regards to the state of the dice as we roll it: if we rotate the dice though 90 degrees around any axis through the centre of two opposite faces, then as far as the mechanics is concerned there is no difference to the dice: this is a symmetry for the mechanics, even though we observe that there is a difference due to the position of the numbers changing). The premise behind the classical interpretation of probability is that every symmetry of the dice and the method of rolling it should be respected in the final probability distribution. What I do when studying quantum field theory day to day is obviously far more complicated than this simple example, but the same underlying principle is still present: symmetries in the underlying model should (baring some well understood exceptions[6]) be respected in the final conclusion. In the case of fundamental physics, this is more reasonable because the symmetries are known (about as well as anything empirical can be known) to be exact. However, when it comes to rolling dice, we know that there is no such thing as the perfectly balanced dice: for all the considerable efforts and skill of dice makers, there is always some imperfection or flaw. And this interpretation offers no consolation when the object under investigation has no underlying symmetry. That the dice is not perfectly balanced is not actually a complete disaster: one can take the actual state of the dice as a small distortion from the ideal state of perfect balance, and, going through the mechanics of dice rolling, incorporate this distortion into the final distribution we derive. Again, we still do not need to roll the dice to derive the probability distribution. However the calculation is very difficult. Except in a few very particular circumstances, either where we know that the symmetries are exact, or we are completely uncertain as to how they are broken (we know the dice is almost certainly unbalanced, but we don't know which number is favoured or by how much, so, since it is just as likely to be one number favoured as another, we still treat each possible outcome as equally likely), the classical interpretation of probability cannot be used as basis of probability. However, its underlying premise, that probabilities reflect underlying symmetries of the system, has a role to play. If there is known to be an applicable symmetry, then we can use that to establish a probability distribution. Indeed, this is the only way to construct a probability distribution which can objectively apply to the real world (which, as a scientist, ought to be our goal). But we need some other approach to extend this understanding to cases where there is no symmetry. Thus the ideas about symmetry on which the classical interpretation is based are a necessary part of the understanding of probability; but they are not by themselves enough. The classical interpretation fails because it claims that these premises are both necessary and sufficient. In practice, they are necessary, but not sufficient.

(2) **The frequentest interpretation:** The frequentest interpretation takes the definition produced at the start of this section most literally. It states that probability

[6]The exceptions are known as spontaneous symmetry breaking.

is the measure of the ratio of the outcomes after we have performed the experiment a large number of times. It states that we cannot in practice deduce the probability distribution, but only measure it. We can then use these measurements to get some idea of what will happen if we repeat the experiment again.

What if we only make a measurement a large number, N, times? The common statement is that the distribution we measure will differ from the distribution for an infinite number of measurements by an amount proportional to the inverse square root of N. Thus the more samples we measure, the closer we are to the true distribution. More accurately, we should think about the distribution of distributions. Each time we take N samples of a particular quantity we get some distribution; for example a certain number of those N samples will record the result 0, a certain number will record the result 1, and so on, and a distribution D is defined as the set of mappings between each possible result and the number of times it is achieved in our sample. We can also repeat this sampling procedure numerous times (say M times), so instead of just one sample distribution D, we have numerous different distributions D_1, D_2, D_3 and so on. We can classify each of these into different types, and count how many distributions of each type we have: this is what I mean by a distribution (X) of distributions (D). If M is large enough, then the distribution X will be sharply peaked (it will be very close to a normal distribution), with a width proportional to the inverse square root of N. So the formal statement is that 95% of the distributions D we measure will be within two standard deviations (2σ) of the true infinite N distribution, with the standard deviation scaling with the inverse square root of N. Now, of course, usually we just measure one D, so this is often interpreted to mean that there is a 95% chance that the results we measure are within two standard deviations of the true distribution; but this statement is over-simplified and misleading. What is meant by a 95% chance in the context of the frequentest interpretation? In reality, we have to go back to the picture of an infinite number of sample distributions Ds, and the 95% only makes sense in the context of this picture. Each sample we take differs from the true distribution by a certain amount, and we cannot say on the basis of the distribution we measure what that amount is; the most we can say is that it either lies within that 2σ bound or it doesn't. Thus the frequentest interpretation in practice means that we either need to make an infinite number of measurements, or take an infinite number of samples with a finite number of measurements to gain the correct distribution, both of which are impossible.

It is understandable that frequency distributions are confused with probability distributions the frequency distribution divided by the total number of observations satisfies the defining characteristics of a probability. So what is the difference between the two? The most important one is that frequency is something we measure, while probability is calculated from a theory or a model. Frequentists, who confuse the two, are effectively robbing themselves of this distinction, which is unfortunate, for there is then no word to describe a theoretical concept that satisfies the defining attributes of a probability, and no means to therefore test the premises of a model by comparing the resultant probability distribution against a frequency distribution. Another related problem is the confusion between models and reality; a model is not the same as reality, and by maintaining the clear distinction between model based probability and reality based frequency we keep ourselves safer from making this fallacy.

I do not think that the frequentist interpretation has any merit whatsoever. There are well known problems with it: how do we assign probabilities to truly unique events? How come the distribution depends on how many times we perform the measurement, and why do we get different answers each time we try to

measure it? But mostly, it just reverses the order of the construction. Once an event has been observed, it is known with certainty what had happened. The probability of a measurement is either 1 or 0. The frequentest theory thus starts with a series of events which are known with certainty, and tries to derive from these something which is uncertain (given that there is not infinite statistics). However, in a logical argument, if the premises are certain and the method deterministic then the conclusion ought to be certain as well. What is measured then is not the true probability distribution, but something else; and that something else is then treated as though it were the true probability distribution. The probability distribution ought to be the premise, and the relative outcomes of the experiment the conclusion. Instead, the frequentest interpretation takes the relative outcomes to be the premise, and the distribution the conclusion. We should use probabilities to predict experimental outcomes given some underlying theory; instead the frequentest tries to determine probabilities from experiment.

(3) **The Bayesian interpretation:** Bayes regarded probability as a representation of uncertainty. Take, once again, the example of rolling the dice. This is a classical mechanics system: in principle, if we know the starting conditions (the initial orientation of the dice, the precise motion we use to throw it, the state of the atmosphere, friction with the table, and so on), then we ought to be able to precisely predict the final result. The reason we can't make this prediction is that we do not know these starting conditions precisely enough, and cannot perform the calculation precisely enough. In other words, our uncertainty in the final outcome is related to our uncertainty in the initial premises. In this way, probability is just a classification of uncertainty. In principle (though in practice this is difficult) we start by assigning distributions to each of the premises (such as how the die is weighted) based on our uncertain knowledge of them, grind the whole process through the process of classical mechanics, and finish with another probability distribution. We take each possible starting position, calculate the outcome, multiply it by the probability of the initial conditions, sum it all up, and that gives us the final probability distribution.

So for every proposition, we label it with 1 if we are certain that it is true, 0 if we are certain that it is false, and the numbers in between represent degrees of uncertainty. Probability is thus only applicable to propositions. More precisely, we can define the set X of every possible mutually exclusive possibility, and suppose that the sum rule of probability ($P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$) applies for any subsets A and B of X (A and B are, like X, sets of propositions). Since the set X is complete, the total probability must sum to one, since we are certain that one of the propositions in X must be true. This is enough to uniquely define how uncertainty is characterised between 0 and 1 for each of the propositions, and it also gives us all three of the defining attributes of probability.

Bayesian probability is of particular importance to scientists: it is the tool we use to classify experimental error. When we measure a quantity, we don't get or conclude a precise result. Rather, we publish that there is a certain probability (usually 68% or 95%) that the final result lies within a certain confidence interval, with a certain central value. We ought to characterise what the probability distribution is as precisely as we can. Now clearly, what we are talking about here is a measure of uncertainty in the outcome of the experiment. We have performed the experiment as best we can, but we still don't know exactly what the result is. But here is a probability distribution which is the best we can offer for the final result. This is just using probability as a measure of uncertainty. Moreover, the overall uncertainty comes from numerous different sources. These are combined using the tools developed in Bayesian statistics to give a final result.

But what does that 95% chance mean in the context of a Bayesian model? It represents an uncertainty, but how certain are we about this uncertainty? When rolling the dice, we computed the uncertainty of the final result by applying classical mechanics to our statements of uncertainty of the initial conditions, but where do we get those first numbers from? They are either arbitrary, or emerge from some other set of premises via another calculation, and so on until we reach an impossible infinite series. In practice, the 95% is derived on the basis of various assumptions and models. For example, we might assume that if we repeat our measurement, then our individual results will distributed according to a normal (or some other) distribution, or that if the resolution of our instrument is 1mm, then our measurements are uniformly distributed within half a millimetre of the central value. These assumptions might well be reasonable, but they are still assumptions and our final result will depend on them. Fine, we should state that model dependence when we publish our results; if possible try out different models and compare what results they give. As can be seen, a good scientist will spend more effort over his error bars than in getting the central value which everybody quotes. But to do so, he assumes the Bayesian interpretation of statistics.

One problem with the Bayesian approach to probability is its subjectivity. Suppose we are tossing a coin. Person A knows nothing about the coin, and how it is biased, so assigns a probability of $1/2$ to the toss being a head, since from his perspective there is no reason to prefer one outcome over another. Person B, however, has examined the coin, and determined that it is weighted so that it will come up heads three times out of four. So he assigns a probability $\frac{3}{4}$. Each of them are perfectly correct in their arguments. So what is the probability of tossing the coin, $\frac{1}{2}$ or $\frac{3}{4}$? Probability has become subjective, which makes it a poor tool to make comparisons against objective reality.

I should emphasise that although (if the Bayesian interpretation is correct) every statement of probability is an expression of uncertainty, not every uncertainty can or should be expressed as a probability. This is most obvious for those things which cannot be expressed numerically or subdivided into clear discrete subsets; in this case it is impossible to use mathematics and therefore neither can we use probability. In these circumstances, our uncertainty still needs to be expressed, but it needs to be expressed in a different way. For example, happiness cannot be measured. Sure, we can ask people to mark how happy they feel on a scale of one to ten, but this does not measure happiness, but only numbers selected by some subjective whim. We can scan the brain and measure its chemistry, but again, this would not be a measurement of happiness but only neural activity and hormonal levels in certain areas of the brain. It might be related in some way to happiness, but unless that relationship is known (which it can't be unless we have a prior numerical measurement of happiness) it is still useless as a means of quantifying happiness. This doesn't mean that happiness can't be studied, but only that it can't be studied scientifically in a rigorous way.

But there may also be other quantities which can be expressed numerically, but for which probability is not the best way to express our uncertainty. The definition of a probability, $P(A \cup B) = P(A) + P(B)$, goes beyond what we need as an expression of uncertainty. In particular, uncertainty in quantum physics is best expressed in a way where this condition is not satisfied. It will lead to a mathematical description of uncertainty, but one which violates this condition and has other conditions in its place.

(4) **Probability as the logic of uncertainty:** Related to Bayes' interpretation, but avoiding the problem I mentioned, is the following concept: probability is the extension of logic which applies in situations where the premises are uncertain, with

the uncertainty of the premises classified using a function normalised to one. The conclusion will also only be stated in terms of a similar probability distribution. Now, it should be emphasised that in this interpretation, probabilities are always model dependent. For example, when rolling the dice, we assumed that classical mechanics was the correct model to describe the dice rolling. The final result we obtain will depend on this assumption. More importantly, we also assume a particular distribution of the starting conditions. Thus the most we can assert is 'The probability of the final result given these premises and given this model is so and so.' The classical interpretation is thus an example of this, where the initial premises (unstated by the classical philosopher of probability; put firmly in place by the one who holds to the logical interpretation) are that the system under study has certain symmetries. More strictly, the premises are that our knowledge of the system has those symmetries; i.e. even if the dice isn't balanced, we don't know in which way it is biased, with each possibility equally likely, so as far as our knowledge is concerned the symmetry remains even if it isn't actually present in the die.

This avoids the subjectivity problem of naive Bayesian logic. Because probability is now conditional on its premises, the statements 'The probability of getting heads if we don't know how the coin is weighted is $\frac{1}{2}$' and 'The probability of getting heads if we know the coin is weighted in a particular way is $\frac{3}{4}$' are now something which everyone can agree on, regardless of their knowledge or ignorance about the coin. The person who knows about the coin will agree that given the premise of ignorance, the correct probability is $\frac{1}{2}$; the person who is ignorant will agree that given a different set of premises, the probability might be calculated as being different from $\frac{1}{2}$. Thus probability is observer (and observation) independent, and therefore objective. The question then concerns comparisons between those premises and the real world to determine which is in accordance with reality.

(5) **The quantum interpretation:** Finally, we can interpret probability as an extension of logic where there is uncertainty in both the initial premises, and the method used to combine the premises to get the conclusion. Once again, the probability distribution measures our uncertainty. We input various probability distributions, and the outcome of our calculation is also a probability distribution, which describes the frequency of our observations if we perform the experiment enough (as many as possible) times. This is clearly related to the logical interpretation; it differs in that the analysis as well as the premises contain uncertainty. This is the interpretation which, I think, best fits quantum mechanics and quantum field theory and other indeterminate systems.

For example a classical mechanics system ought to be treated according to logical probability. Here the initial conditions are not known perfectly, because they cannot in practice be measured to such accuracy. Instead, they are known to within a certain degree of uncertainty, and from this the conclusion is also computed as a probability distribution. However, the means to get from the initial state to the final state is fully determinate. In a quantum mechanical system, the initial state is again not known to complete accuracy, not only because in practice the errors of any measurement preclude certainty, but also even in principle because it is impossible to simultaneously measure certain distinct observables. But even that is not the whole source of the uncertainty because the path of the particle is indeterminate. In classical mechanics, a particle can travel from A to B only by one fixed path. In quantum physics, it can take any possible path to get from A to B, each path weighted by a certain probability amplitude. Finally, we have further indeterminacy when we take the measurement.

There is a second difference in the quantum mechanical version of probability. Uncertainty in quantum physics is not at the most fundamental level expressed in a manner consistent with the axioms of probability. We can convert this expression to a probability, but we cannot go the other way from a probability to an unique quantum amplitude. Probability is a natural way of expressing uncertainty because of its link to frequency distributions. This link is particularly important in experiment, because you obtain a frequency distribution whenever you count numbers of different things, and counting is one thing that experimentalists are particularly good at. But we should not be confused by this by saying that probability is the only way of expressing our uncertainty. For example, we can think of representing uncertainty for a particular proposition (for example 'A particle is located at a position x) by a circle of unit radius centred at a point O. The centre of that circle implies that we are certain that the proposition is true; the circle's circumference represents certainty that the proposition is false. This expression is natural if there are many ways in which a proposition can be false, but only one way in which it is true. For example, if we ask 'Is that bus red?' then if the statement is true then it is of one particular colour, but if it is false, the bus could be yellow or green or white and so on. Not every non-red bus is of the same colour, and a simple linear scale of uncertainty lumps together objects which are in practice distinct. Alternatively, if there are many ways in which a proposition can be true, but only one way in which it can be false, then we can invert this description and have the perimeter of the circle representing certainty of truth, and its centre certainty of falsity, and the middle as varying degrees of uncertainty. For example, if we ask if an electron is present at a particular point in space and time, there are numerous different and distinguished states which that electron could be in, so a single point like expression of uncertainty is not sufficient to describe our knowledge and yet also distinguish between the different states. However, there is only one way in which the electron can be absent. This circle notation is thus natural for particles which could be present, but in a number of different ways which need to be kept distinct.

This representation is not a probability, since the position on the circle is not guaranteed to be positive, however (as this is used in physics) we can derive the probability from it. If we select one point within the circle as representing our level of certainty concerning some proposition, the probability is proportional to the square of the radius at that point.

Now it might seem that this representation is a bit redundant – we have numerous points on the circle representing the same probability. And it is true that the system has a rotational symmetry. In the physical application, this rotational symmetry is hugely important – it directly implies the existence of the electric and magnetic forces. But the redundancy is only unnatural because we naturally think of uncertainty in terms of probability rather than this circle picture. Two points with the same radius on the circle are not equivalent; they just happen to give the same probabilities.

For example, if we take the average of two large probabilities, we get a large probability. So a high degree of certainty plus a high degree of certainty equals a high degree of certainty, which fits our naive intuitions. This is because all probabilities are positive. However, with the circle picture, if we average between two points on the circle which are close to the circumference, we might get another point close to the circumference (meaning a high degree of certainty that the proposition is true) if those two points are each located at the same end of the circle – again, we get a high degree of certainty plus a high degree of certainty equals a high degree of certainty. But if they are located at opposite ends of

the circle, then the average position is in the middle. Two propositions we have each assigned a large degree of certainty to can together equal one thing that we are certain is wrong. This is known as an interference effect, and it is of huge importance in physics.

So why bother with this intuitive measure of uncertainty? Because that seems to be the measure we need when studying quantum physics. We call this measure of uncertainty an amplitude $\psi(x \setminus A)$ (I will later distinguish amplitudes from probabilities by using the words *likely* or *likelihood* to indicate amplitudes and *probable* or *probability* to indicate probabilities), where the x reminds us that we need a different circle at each location in space. As with the previous interpretation of probability, we can compute ψ based on various premises (this interpretation of probability still represents an extension of logic, so we must state the premises along with the conclusion), denoted by the $\setminus A$; since we are describing a physical system, A will include some initial set-up of the system (perhaps expressed with perfect certainty, perhaps with some uncertainty), and will also depend on the laws of causality that describe the evolution of the physical system over time. The laws of causality that govern quantum mechanics and quantum field theory require those seemingly redundant degrees of freedom given by the amplitude - the laws act on states described by amplitudes rather than probabilities. There are reasons why this is the case (though this is not the place to explain those reasons). Probability is too crude a measure to capture the internal symmetries of an electron. Therefore we use the amplitude, which does allow us to reflect the symmetries.

So in quantum physics, although each particle follows a particular path and can be found in a particular place, we cannot be certain of where that is even with perfect knowledge of the initial state of the system because causality in quantum physics is indeterminate. We represent that uncertainty not in terms of probabilities, but in terms of these amplitudes. We know the likelihood of a particle moving from one state to another (from the laws of quantum mechanical causality), with each state represented by a point on the circle of uncertainty, and this leads to a prescription of how the amplitudes change in time (in quantum mechanics, this change in the amplitude is deterministic; in quantum field theory it isn't). This, of course, is just the same as a process of logic where we move step by step from premise to conclusion. So a quantum amplitude is a statement of uncertainty whose squared radius is a probability (and thus with the defining characteristics of probabilities) and which represents an extension of logic, being calculable from various premises (thus it is an objective rather than subjective uncertainty). In the process of measurement, the wavefunction collapses because there is no longer uncertainty over the measured properties of the particle.

There are other less common interpretations of probability, such as probabilistic observations arise from measuring apparatus rather than being inherent in reality, but these in general don't answer the question of what probability fundamentally is and thus I can safely neglect them.

Probability is therefore best seen as an extension of logic (Briggs, 2016), especially for the applications to theoretical physics where I need it. It is a means to reach uncertain conclusions from uncertain premises or an uncertain algorithm. However, the uncertainty is characterised and obeys certain well defined rules. In this way, the results of this reasoning are not random, but still (in some senses) predictable. Causality (according to the classical definition), and the process of explanation, is still maintained. Also note that if probability is an extension of logic, and every logical argument is conditional on its premises being true, then this means that all probability is conditional. We cannot say that the probability $P(A) = \frac{1}{2}$; we can only say that $P(A) = \frac{1}{2}$ if assumptions $X, Y, Z, \ldots$ are satisfied. The

probability of a coin toss coming up heads is $\frac{1}{2}$ if the coin and the method of tossing is perfectly symmetric; if the coin is not perfectly symmetric we will come up with a different number which will depend on the various assumptions we put in. A probability itself is still regarded as a measure of uncertainty, with 1 meaning that we are certain the proposition is true; 0 meaning that we are certain the proposition is false; and subject to the addition, normalisation and positivity rules discussed above. The power of probability to predict the distribution of (infinite) samples of data thus allows us to test of whether those assumptions each probability depends on are in correspondence with reality.

One objection to the logical interpretation of probability is the status of unproved mathematical axioms, such as Goldbach's conjecture, that every even number greater than true is the sum of two prime numbers. We don't know whether this is true or false, since it hasn't been proven one way or the other, so on an an interpretation where probability is a measure of certainty (which includes the logical interpretation) we should be able to assign a certain probability that isn't one or zero to it being true. But the truth or falsity of Goldbach's conjecture would follow necessarily from the underlying axioms. So on the logical interpretation it must be either true or false, and therefore have a probability of either one or zero. This contradiction is used to argue against the logical interpretation. But it makes an error. In the logical interpretation, every statement of probability is conditional. So asking what the probability of the conjecture being true is meaningless. Instead, we should ask what the probability is conditional upon various assumptions. So in the first case, we have the probability that Goldbach's conjecture is true given the standard axioms of mathematics, and this will be either 0 or 1 but we don't know what because nobody has been able to do the calculation. To convert that uncertainty into a probability expressed as a number, however, we need further axioms, which will exclude the axioms of mathematics (otherwise those alone would make the probability 0 or 1), but will include such premises as nobody has been able to find a counter example, and human skill at disproving mathematical conjectures of this sort. So we are comparing the probability of Goldbach's conjecture being true given one set of conditions against the probability of Goldbach's conjecture being true given a different set of conditions. Since these are two different things, there is no contradiction in saying that they lead to different numbers.

Implicit in this response is the idea that not every expression of uncertainty can or ought to be expressed as a numerical probability. Sometimes we simply can't know (yet) how to calculate the answer is. Some things cannot be captured by a numerical scale. And that's fine. In those cases we simply don't use probability to express our uncertainty, but either turn to other methods or simply remain agnostic until further evidence comes in. It is perfectly acceptable acknowledge imprecision. To say that the answer is either 1 or 0, but we don't know which, so we will leave it at that until further evidence comes in, and, if the result has some practical effect on our lives, prepare as best we can for each possibility.

I should also mention statistics, and in particular statistical modelling. This is something which is widely used, but which we should be cautious about. A statistical study does not measure causation, but only correlation; and worse, not correlation in nature, but only correlations in a limited data set. The statistical scientist will parametrise whatever they are studying in terms of some numerical quantity, perform various tests, and say something about one of those parameters, or how one of those parameters varies with another parameter if some particular fitting ansatz is chosen. But all this does is state something about the data set. It does not state whether or not the data set models reality, or whether the parameters accurately model something important in reality. They might well do, of course; if the study is performed well then they would to a certain extent. But we must never interpret the numbers in a model as though they were reality; all they are are certain parameters of a simplified model (which often has no theoretical justification beyond that it happens to fit this particular data set reasonably well; and of course there could be numerous models leading to dramatically different conclusions which fit the data well) and they should be

treated as such. But, most importantly, such a study cannot tell us why the data correlates as it does. It cannot tell us what causal effects are present, or if any are present. Is X causing Y or Y causing X or are they both caused by some unmeasured quantity Z or is the correlation between them just a coincidence? Statistical analysis cannot tell us. We can only know about causation through theory, and not through direct observation. Statistical models, which are derived from models fitted to data, are not theoretical models, which are a priori but then compared against data.

2.3. The Scientific Method

The scientific method is a means of applying logical analysis, and in particular the mathematical branch of logic, to observable systems. There are, in fact, two principal tools used by the scientist. The first is categorisation: the separation of beings into certain well defined types. For example, we might think of lists of biological species, or the enumeration of the chemical elements, or the list of the fundamental particles in physics. The procedure in question is basically just to study the beings in question, note similarities, and put them into categories. Of course, the tools by which we note the similarities and differences may change and grow more sophisticated over time: it used to be that physical similarities were the main driving force behind biological classification; now more weight is put into genetic studies. But the basic process is the same: to form a list of certain objects in the real world, and describe their properties, similarities and differences as well as we can.

The second tool is of more interest to me in this work. It is to model aspects of reality mathematically and, in particular, to try to be able to understand how various bits of matter interact with each other, so that it is possible to predict future behaviour and make use of our knowledge for technological progress. For me, the main interest is one of curiosity: I want to understand. For others, particularly non-scientists, this understanding is just a step in applying the knowledge for the general benefit of mankind. The means by which we do so is to firstly extract certain aspects of reality in terms of mathematical abstractions: perhaps numbers (such as locations in space-time), perhaps as groups or aspects found in topology or differential geometry. A mathematical model is then constructed to describe how these quantities evolve over time. Then, when we have reached our target time and place, we convert back from the mathematics to make predictions about the reality. The most important thing to note is that the mathematical model is not reality: it is at best a description of certain aspects of reality. Through the mathematics, we have come to predict the existence of certain aspects of reality (for example new particles) that are otherwise unobservable. But again, these predictions do not make the mathematics real. While acting as a physicist, I can tell you what an electron will do in various conditions. I cannot tell you what an electron fundamentally is. Physics, therefore, cannot contain all of the possible knowledge.

Now it should be clear from this that science is only useful for those observables and those areas of study which can be described by mathematical objects. Many parts of human knowledge cannot be so described, for example ethics. Human happiness or well being or virtue cannot be characterised as a number. This does not mean that it cannot be studied logically or rationally. It means that other tools are best suited to address the problems. To be unscientific does not necessarily mean that it is wrong, illogical, unintelligible or irrational. It is always best to use the appropriate tool for the job. The scientific method is successful at modelling certain aspects of the physical world, and through that certain aspects of chemistry and biology and other parts of knowledge built upon those subjects. It is not successful when applied to fields wholly unsuited for it.

The mathematical science is built on two legs, theoretical models and experiment. The model is ultimately what we want to uncover; experiment or observation our means of doing so. The claim is that there is a single true model which describes all of those aspects of reality which can be modelled mathematically. If we are lucky, this might be expressed

in a single equation. I personally do not think that we will be so lucky, but nonetheless we will hopefully uncover some simple set of principles which can be applied universally. We are trying to discover what this true model is, and subsequently find the full range of consequences of it. Our current best theory is not this true model. However, we can hope that we are fairly close to it, and over time the differences between our best theory and the model which truly represents nature are gradually eroded.

Now, it should already be observed that certain philosophical presumptions have already been made before we can even start building a scientific model. I list a few of them below:

- There is a real world out there.
- This real world is objective; independent of any human observer – it does not matter who performs an experiment.
- This real world is at least partially describable using a mathematical representation.
- The real world is in principle predictable, either probabilistically or deterministically (this leads to the notion of causality).
- There is change in the real world (since change is what we are interested in measuring).
- There is a single model which correctly describes those aspects of the real world which can be reduced to mathematics.
- This model is intelligible, at least in principle (if not necessarily in practice).
- There are certain universals in the world, simple building blocks of matter which each behave in a consistent way no matter where they are in the universe.
- We can make reliable observations concerning this real world.

Some philosophers have denied some of these assumptions. I cannot see how any scientist can consistently do so; for example, I cannot see how any scientist can consistently be an idealist or nominalist.

So how does the scientific community proceed? It is possible to summarise the process simply (at least for fundamental physics; other branches of science might modify this method as is appropriate to their area of study): we first of all make a guess of the model. If we are fortunate then we will have many different guesses. We then derive predictions from this guess, ideally for experiments not yet made, but also for experiments already performed. We perform the experiment. If the predictions are accurate, then we make more predictions, until we find one that fails. If the prediction fails, then we either modify the theory appropriately, if that is possible, or abandon it. In this way, we systematically eliminate incorrect guesses until we are left with only the true theory. The scientific method is thus a combination of the logical steps of deduction and comparison.

In practice, however, there are other ways barring experiment and the needs of mathematical consistency by which we can assess our theories. In particular, the concept of symmetry has been particularly important for whittling down the plausible guesses in modern physics. For example, when mapping from reality to a mathematical representation, we need to specify a coordinate system; and our theories are expressed in terms of this system. However, these coordinates are not part of reality, but imposed on it by us, and (at least to a limited extent) we might expect that the underlying theory should be independent of our choice of coordinate system. This implies various symmetries. This conclusion is not certain, but it is still an idea which can guide us as we try to think of possible theories. If it is found to be satisfied very well, we can draw it up as a general principle.

So how do we construct the guesses? One means is to use induction: use the results of experiments to guide us. We do not confirm hypothesis via induction, but it can be useful for framing them in the first place. Secondly, we have philosophical considerations. These were particularly (though by no means solely) important in motivating Galileo and Newton and their generation, and also Einstein and his colleagues in the formulation of the theory of relativity. Thirdly, there is intuition. Fourthly, there is good old-fashioned

trial and error: write down as many crazy ideas as you can think of, until you find one that works. Science works at its best when all of these methods of forming hypotheses are used together. In recent years, unfortunately, progress has been slow in particle physics chiefly because experiments have been overtaken by theory and so don't provide the help we need, there has been no good philosophy to provide inspiration, and intuition is misleading in realms that behave very differently from our everyday experience. We have basically been reduced to trial and error, and that is necessarily slow when it is difficult and time consuming to compute the consequences of a theory.

Of course, this method is made more complicated because neither theory nor experiment are perfectly precise. In almost every case, we cannot say 'The answer is this,' but only (for example) 'There is a 95% probability that the answer lies between these values.' Where does the theoretical error come from? Surely theory, as a matter of deduction, should give precise conclusions?

There are several sources of theoretical error. Firstly, we depend on some parameters which can only be measured by experiment and not estimated theoretically (Gödel's theorem strikes again). None of these are known with absolute precision, and this leads to imprecision in any prediction made by the theory. Secondly, we often cannot compute exact results from the theory, but have to rely on some controlled approximation. These approximations are controlled; we can say with some confidence roughly how far we are from the true solution; and we can always improve the approximation by spending a few years computing the next most important contribution to the final result, but we cannot reach the true solution to a perfect precision. Thirdly, we sometimes need to extrapolate our results, either because we cannot perform computations at precisely the right parameters, or because we need to match with some notational schemes used by the experimentalists but which aren't so straight-forward theoretically. Every extrapolation brings in certain errors: it requires some fit, which always increases the error, but more importantly extrapolation only makes sense in the context of a particular model. Which model should we use to control the extrapolation? Are we using the model correctly? This is another source of uncertainty. Finally, it is sometimes necessary to use numerical simulations to extract observables from the theory. These are generally statistical, and converge to the right answer given enough data, but usually we don't have enough data to get there. This leads to a further source of imprecision.

The key thing is that all of these sources of error can be estimated and controlled. In principle, we could reduce them towards zero – it just requires a lot more effort than we are capable of putting into the problem. In practice, over time, we do reduce them towards zero: much of science is about reducing these uncertainties by as much as we can. But nonetheless, we can obtain some estimate of how far away we are from the true result. Scientific uncertainty is not a free-for-all, anything is possible type thing. It is kept within certain carefully measured and increasingly (over time) small bounds.

So what of experimental error? Observations ultimately come from our senses, sight, taste, hearing, touch, smell, and so on. However, to perform a scientific measurement, we almost always need to use some instrument to help make our senses more precise, or to more accurately quantify what we see. For example, to measure length, one method is just to look at the object and make a judgement. But we don't see length, we see objects. Length is just some artificial construction we apply to objects. There is therefore a mental conversion from what we see to our estimate of its size, and this mental conversion is not usually very precise. The difference between a scientist and a casual observer, however, is that the scientist is fully aware of that imprecision and tries to accurately state how large it is. Another method, more precise, is to use a ruler or other measuring stick. More precise still would be, if possible, to use an interferometer. In each case, we use certain tools to improve the natural abilities of our eyes. Obviously, each additional tool we use carries with it another possible source of error: is this tool trustworthy? Are the marks on the ruler made at the correct points? Are the mirrors in the interferometer correctly aligned, or the

wavelength of the light being used accurately measured? Usually, however, the gain we get from the inherent accuracy of the tool more than offsets the additional errors coming from it. Again, we need to carefully quantify and state each of these uncertainties. These are known as measurement errors.

But whatever tool we use, ultimately we require using our eyes. This begs the question: are our eyes reliable? We know that there are occasions when they are not: when we are drunk or under the influence of some other narcotic, some optical illusions take advantage of imperfections in how the data is processed in our brain (or elsewhere in our optical system), we have difficulty in resolving precise details at very large or very short distances. Some have tried to use such limitations to cast doubt on our eyes as in any way reliable. This, of course, is foolish. That our eyes are limited in some circumstances does not mean that they are not usually reliable. Instead, we should bear these limitations in mind when we make an observation: we know under which circumstances they apply, can often quantify how serious they are, and can take measures to minimise them, and, if they cannot be eliminated, to include any such uncertainty in the final statement of our result.

It is usually best to make numerous observations of the observable, to check our result if nothing else. Each time we make the observation, we most likely will get a slightly different result. How then do we extract a final number? The easiest way is to take the mean (or average) value, perhaps after excluding any points where our equipment was misaligned or some other serious and known problem. However, the mean value is not going to be the same as the true central value of the distribution of measurements , unless we take an infinite number of measurements, which is impossible. Therefore this process will lead to another uncertainty, the statistical error. However, once again, this can be estimated, and stated alongside the result.

Then we often require theory to interpret the experiment. It is rare these days that we directly measure the observable that we want to study. Instead, we use some device which measures something else (and perhaps displays the result on a computer screen), from which we can deduce the value of something else, from which eventually we reach the quantity that will be quoted in the final paper. Theory (some previously well established theory) is required to firstly understand what precisely the measuring devices are in fact measuring, and secondly to understand the chain that links the display on the screen with the quantity that we desire to measure. Again, no theory is perfectly precise or can be perfectly trusted (given that we can't be completely certain that the theory is right): another source of imprecision. Again, this imprecision does not mean that the final result is inaccurate: we can quantify the uncertainty as a probability distribution (in the Bayesian interpretation of probability), and restrict it to within certain bounds which can be computed. But it still needs to be taken into account.

So when we compare theory against experiment, we are not comparing two numbers, but two narrow bands, which progressively get narrower as we improve the calculations and the experiments. This means that falsification is not just a simple means of taking one experiment and comparing it against one theoretical number: rather, it is a matter of measuring the overlap between two probability distributions, and thus any comparison can only be expressed in terms of a probability itself. It can take a great deal of effort to firstly accurately quantify the imprecision and secondly reduce it to the level where we can truly distinguish any inconsistencies between theory and experiment. Every scientific measurement or prediction for a continuous observable includes an estimate of the uncertainty; if there is no statement of the uncertainty and the degree to which it depends on various models and assumptions then the measurement is not scientific.

Finally, I should note that the process of falsification can make mistakes, simply because theory (which could be wrong) is required to interpret experiment. Two examples will suffice. The story of Galileo is well known: persistent defender of Copernicus, ridiculed and persecuted by his contemporaries, but who eventually rose triumphant when Newton

vindicated him as the founder of modern science. What is less well known is why he was opposed by his contemporaries. The main reason was that Copernicus' system was (so it seemed) firmly 'disproved' by the best experimental evidence. If the earth rotates around the sun against a fixed background of the stars, then there is a slight shift in the position of the stars near the poles (the North Pole star being visible in Europe) depending on the time of the year. The slight difference between the earth's position in summer and winter means that the light from the star comes to us at slightly different angles. This is known as the stellar parallax. Those of Galileo's era calculated the distance of the stars by assuming that they were roughly the same size as the sun (good enough for these purposes), using their apparent diameter in comparison to the sun's diameter and the known distance from the earth to the sun to estimate how far they were from us, calculating what the angle of parallax should be, and measuring it. While the measurements were obviously less precise than later measurements would be, they were nonetheless good enough to distinguish the angle they had calculated from zero. No parallax was found. Thus, it was stated, the earth was stationary against the background of the stars, and thus Copernicus' system (and Kepler's better system, though this was ignored in Italy at the time of Galileo by both sides in the debate) was definitively ruled out. Galileo himself had no answer to this to the day he died: he could only appeal to the simplicity and elegance of the heliocentric model, and was eventually forced to argue for his views from theology (which brought the Church into the debate). The answer to this problem would not be known for certain for well over a century (although diffraction was first observed in the late seventeenth century): diffraction and atmospheric distortion make the stars appear larger than they actually are; the astronomers of Galileo's time had underestimated their distance from us by a factor of a thousand, and consequently the size of the parallax by about the same amount. The actual value was beyond the resolution of their measurements. It was not until the eighteenth century that the solution to the parallax problem was proposed[7], and the early nineteenth before it was measured[8].

My second example is more recent. In the early nineteen fifties, two young scientists, Chen Ning Yang and Robert Mills, developed the formalism of non-Abelian quantum field theories, which they proposed to explain the weak and strong nuclear forces in a similar way to the then rapidly becoming accepted theory of electromagnetism. Before publication, Yang presented his ideas to a seminar at Princeton, at which the distinguished physicist Wolfgang Pauli happened to be present. Pauli was one of the world's best physicists, and among the most important people involved in the development in quantum mechanics in the early twentieth century; Yang at that time still relatively unknown[9]. Pauli, who had been working on similar ideas, knew the flaw in Yang's theory: to be consistent, it needed (in its simplest form) three massless particles or (in its next simplest form) eight massless particles. These particles should have been observable, and had not been observed. Yang and Mill's theory was unquestionably elegant, but it was also unquestionably wrong. Yang had no answer to this question on the day to Pauli, and he had no answer when his paper was published. Fortunately, the referee still let the idea pass. The particle physics community concentrated on other ideas, and the Yang-Mills theory was almost entirely ignored, though not quite completely ignored, since a few people still followed it up. But, gradually, the alternative ideas started falling away, while the problems with Yang's theory were solved. In the nineteen sixties, Brout, Englert, Higgs, and a few other scholars showed how the gauge Bosons mediating the weak interaction could acquire a mass through a phenomenon

[7]Bradley in 1729, over eighty years after Galileo's death; it was some time before this was accepted.

[8]Callendrelli in 1803; the Coriolis effect, discovered in 1790, was also crucial evidence for the rotation of the earth.

[9]The thing that Yang was best known for was a 'ludicrous' suggestion that the weak interaction could break parity symmetry, which everybody laughed at since it was obviously absurd. Two years after this encounter with Pauli, Yang's suggestion was confirmed by experiment. The year after that Yang and his co-author, Frank Lee, won the Nobel Prize.

known as spontaneous symmetry breaking. The work of Yang-Mills, Yang-Lee and Higgs were put together by Glashow, Salam and Weinberg to create the now standard electroweak model, unifying the weak nuclear force with electromagnetism. At the end of that decade, and the start of the nineteen seventies, it was realised that the observable states of the strong interaction were not the gluons (Yang's massless gauge Bosons), but bound states which did have a mass (plus corrections to the observable gluon mass from quantum fluctuations). This realisation, and the demonstration that only Yang-Mills theory allowed the observed phenomenon of asymptotic freedom in the strong interaction, earned Gross, Wilczek, and Politzer their Nobel prize. The mathematical consistency of Yang-Mills was shown in the early 1970s, and this earned Gerard 't Hooft his Nobel prize. In the mid 1970s, with experimental evidence for both the strong and electro-weak theories starting to become irrepressibly strong, the standard model of particle physics was established with Yang and Mill's theory as its central element. By the time I was a graduate student a quarter of a century later, Yang-Mills was as established (and as important) a part of the physics curriculum as Newton's, Maxwell's or Einstein's work. Their final vindication came in 2012 with the discovery of the Higgs Boson confirming the ideas of Brout, Englert and Higgs.

These examples (and there are plenty of others) show us that we should be careful with falsification. It may be that the theory used to interpret the 'falsifying' experimental evidence was wrong. It may be that the idea was not wrong, but just incomplete or needs only a small modification. A theory which was previously 'falsified' should become open again if the ideas that seemingly disproved it are later overthrown. It may, of course, be immediately overthrown again when the data is reinterpreted in the light of the new established theory – and most of the time it will be – but that should be shown, not assumed. It is sometimes worth taking another look at old ideas. Of course, these occasions when the methodology of falsification break down are uncommon. The method has been exceptionally good at weeding out both crackpot ideas, and ideas which seem perfectly reasonable (and almost unanimously agreed) until they have an unfortunate confrontation with reality. If something runs against experiment, it may still be worth examining it in greater detail in case those problems can be overcome, but it should be considered rejected until and unless those problems are resolved. Of course, that does not excuse the persecution that was directed against Galileo.

But nonetheless, the scientific method: deduction, comparison, and elimination of incorrect answers, has proved to be very successful in those areas where it is applicable. We have got a theory which is tested to an exceptionally high degree of accuracy, and found to give answers in agreement with experiment. In fact we have got two theories describing different areas of physics – one the very big, and the other the very small, which overlap in the middle (which is the world of our day to day experience) – and each of them are so good that anybody but physicists or other perfectionists would by now be slapping themselves on the back for a job very well done. These theories are, unfortunately, inconsistent with each other, but you can't have everything.

Physical knowledge is thus based on two legs, experiment and theory, and both are essential. The theory is some abstract concept, expressed mathematically, which describes certain aspects of reality. We should not confuse our model with reality. We should not say that because our model behaves in a certain way, then reality also behaves exactly like the model. There are two reasons for this. Firstly, our model is almost invariably only approximate. For my own field of particle physics, we are probably as close as any scientists in reaching the correct model. But we are not there yet. In other fields, such as climate science, the situation is worse: firstly, it is still a relatively young field compared to particle physics, and therefore not so well explored, and secondly it is considerably more complicated than we see in particle physics. There are a huge number of factors influencing the environment. Some of these we understand well; others not so well, and others we may not have even thought of. Understanding these factors in isolation is not enough, because

they interact with each other in a way that means that if you put the two of them together you get something other than just the sum of their parts. Furthermore, the system is inherently chaotic, making predictions over long periods of time difficult: a small error in your starting position can lead to a large error in your final result. This doesn't mean that we shouldn't try to build models, and use them. Far from it. Building an intellectual model is a necessary step in understanding the system. We can't understand without the theory. But it does mean that in fields such as this we need to be particularly careful in testing our models, and particular careful and honest in expressing the uncertainties, and humble when presenting our results until our theory is established by centuries of successful and precise predictions (for that would be equivalent to the standards used elsewhere in science).

The second reason that we should not confuse reality with the model is, even if we find the true model describing nature, that itself is not reality. We should not think that the model or theory captures the whole of reality. This should be obvious, since the model is just an abstract concept, that we can neither see nor touch but which resides in our intellect, and reality is something tangible, material, that we can see and touch and which is exterior to ourselves.

There is, then an important but too rarely asked question: why does science, and in particular physics, work so well? What does an abstract model have to do with a concrete lump of matter? Many philosophers, starting from Descartes, have separated the world of ideas and the world of matter and placed a chasm between them. This truly leads to the question of why physics works. If the world of our intellect is so separate from the world of matter, then why does matter follow paths that we can understand? There are several possible answers to this. Descartes himself proposed that matter should be reduced to things like extension and mass; the quantities we measure and model. However, this does not answer the question but merely avoids it by denying one half of the problem: one cannot touch and see mere extension, yet we touch and see matter. Others deny the existence or relevance of the material world; some deny the existence or usefulness of the abstract concepts (and say the best we can do is note down certain regularities or constant conjunctions, with there being no reason why matter should act that way in principle); a few express the reality of the both the abstract concepts and the material world, but deny a fixed connection between the two: our theories only resemble nature to an approximation, but in truth they cannot even in principle capture any of it, and matter itself is not bound by any rules. My own opinion is that none of these ideas are tenable. All of them leave the question of why physics works ultimately unanswered: they just try to solve the problem by denying one aspect of it.

There is, however, an alternative: to say that a physical object contains an immaterial aspect as part of it. We don't place a chasm between the material and immaterial. We bring them together. Instead, we say that a physical object is not just the matter that we see and touch, nor just an abstract idea such as extension or mass, but the two of them fused firmly together into one being so that you cannot tell where the one component begins and the other component ends. The abstract model we try to find in physics describes the immaterial part of the being; and obviously there is no problem in describing something immaterial in some abstract language such as mathematics for those parts of the immaterial aspect of a being which are suitable to be treated mathematically. Because the immaterial aspect is an essential part of the being, physics can describe very well many of the ways the entire physical object behaves. Physics works because what we describe in the equations is something that is really present in reality (only the abstraction in our understanding is translated into a mathematical language rather than whatever language nature itself uses). Once we have our theory right, then the abstract concept we ponder in our intellect is the same (or equivalent to, after a process of translation) as the abstract concept as it exists in nature. The only difference is in one case it is instantiated in our minds and in the other

case it is instantiated in a physical object. And this, to my mind, is the best way to account for the success of physics in describing some of reality.

In this work, I will assume that our best current theories are pretty close to being correct: only minor modifications are needed to bring them together. This assumption could, of course, be false - physicists thought that they were close to the final answer at the end of the nineteenth century, and were proven to be dramatically wrong; if so, then this work would obviously at best require heavily modification and possibly to be discarded. However, one has to start somewhere, and our best current understanding is as good place as any to begin.

This is not primarily a work about physics; but a work about philosophy. Now how shall we judge philosophy, and in particular metaphysics, the queen of philosophy? There are three tools at our disposal. The first is whether it is self consistent. The second is the explanatory power it has – how much does it explain? The third is what physics can be deduced from it. For metaphysics is linked to physics. Bad metaphysics is inconsistent with good physics. Of course, the process is difficult, and riddled with uncertainty: we are one step further away from experiment. But we should be able to eliminate certain metaphysical possibilities and assumptions. It seems clear to me that ever since the days of Newton, philosophers have been lagging some distance behind the physicists. Newton, Galileo and their contemporaries were inspired by a philosophical revolution that in part preceded them and in part built upon their work. We know that revolution as the enlightenment. My reading of philosophy is largely that it (with very few exceptions) has not progressed much since that time; it is still stuck very much in an enlightenment, and thus a Newtonian, mindset. People like Kant, Hume, Locke, Spinoza and Descartes are still very much admired, even though they based their work on old and by now discarded physics. But physics has moved on. We now know (with as much certainty that experiment can give us) that Newton's description of reality was, at a fundamental level, completely and utterly wrong. So what does that say about the various philosophical systems which in part inspired and in part were built upon Newton's work? They need to be re-examined in the light of the new physics. Given the assumption that we are close to the final answer in physics (and so won't be thrown aside by another quantum revolution), now is a good time to start doing so. In my view, the enlightenment philosophies need to be almost wholly discarded, since they are built on assumptions which can no longer be regarded as consistent with physical theory. They need to be replaced, either with something new, or with something very old[10].

[10]I do not mean that all recent philosophy is wrong. Firstly, there has been exceptional progress in logic and analysis, and perhaps some other areas, which is untouched by any physics. Secondly, wrong premises do not necessarily mean a wrong conclusion - there could well be much of value in modern philosophy. What I mean is merely that we need to re-examine the understanding from first principles: no longer be satisfied just to cite Kant (for example) as a reliable authority, but first question whether Kant's whole system still stands in the light of the new physics, and then ask what we can salvage from its wreck.

CHAPTER 3

Metaphysics

3.1. Introduction

Why do we need metaphysics? Firstly, what is metaphysics? The word means 'that which is above or beyond physics', or 'that which is more fundamental than physics'. In the natural order of knowledge (going from first principles to final conclusions), first we have logic and mathematics as the most basic, then metaphysics, then physics and some other branches of philosophy which are derived from metaphysics and either mathematics or logic; then other branches of the natural sciences and philosophy build on these; and finally we have practical subjects such as politics, history, engineering, and economics which are built on the natural sciences, philosophy, and perhaps some other assumptions and methods unique to each field. Mathematics and logic can be derived from various axioms which we take to be self-evident; physics we learn about from experiment and direct observation of nature, and we construct physical theories through experiment and mathematical reasoning. Metaphysics fills the gap between the self-evident logic and the experimentally derived physical sciences. It also provides the link between physics (and hence observation) and more practical areas of philosophy such as ethics or theology. In a well thought out, complete, and consistent model of knowledge metaphysics gives rise to the fundamental axioms of physics, the fundamental axioms of ethics and the basis of our theoretical knowledge of God. In principle, we use induction (and a fair helping of common sense) to list various possible metaphysical systems from our understanding of physics, and use those metaphysical systems to deduce the premises of ethics, or theology, or politics, or whatever subject we are particularly interested in. Of course, that is the principle; in practice it is one of those methods which is much easier to describe than to do.

The traditional definition of metaphysics, which you will find in all the textbooks (Loux, 2006; Lowe, 2002; Mumford, 2012; Feser, 2014; Aristotle, c. 330BC; Aquinas, c. 1272), is this: if physics is the study of different actual beings, metaphysics is the study of being in itself: what does it mean to exist? Why do certain things exist, and others not? Why do things interact in the way they do? And so on.

I would phrase it slightly differently: if you give a competent theoretical physicist a list of symmetries, and the types of matter that exist and how they are transformed under those symmetries, and whether the system should be classical or quantum; then the physicist will be able to predict how that matter will behave, how it will be organised and manifest itself, how it interacts with other matter, what observable properties it will have, and so on. He will find the most general theory consistent with those symmetry principles, and use it to make predictions about nature (these predictions may be certainties – up to the theoretical imprecision – if the theory is deterministic, or incorporate probabilities if the theory is not deterministic, but both are predictions). The experimental physicist can take those predictions, and compare them against the observed world. Between them, they can find out what symmetries exist in practice, what values the various parameters not fixed by symmetry take, and so on (and, of course, subsequently use that information for more practical purposes to enrich our lives). The physicist cannot, however, at least by using the techniques and methods of physics, explain why the world has those symmetries; why it has those types of matter fields and not others; why there is any matter at all; why

the parameters take those values and not others; why physical law is quantum or classical; why certain interactions are possible and others forbidden. That is, in part, the task of metaphysics: to justify the premises used by the physicists, and derive them from more general principles. The physicist explains what is; the metaphysicist attempts to explain from first principles why it is. Logic is the start of philosophy; but metaphysics comes next, and everything else depends on these two.

Now it is immediately obvious that metaphysics is much harder than physics, at least if we want to do it well. Firstly, because the metaphysicist needs a solid grounding in the correct laws of fundamental physics. This is a particular problem since we do not yet as a species know the true laws of physics, and it is even more of a concern for those philosophers who lack a specialist training in physics. Even for physicists themselves, it is only relatively recently that it has been realised that the methodology I described in above paragraph is how we should be working: the fundamental importance of symmetry was only gradually realised as the middle two quarters of the twentieth century unfolded. The competent physicist can be content with just understanding physics (and as much applied mathematics that he needs); the truly competent metaphysicist must be expert in both physics and philosophy, unless he has no desire to test his ideas against reality. The second reason why to do metaphysics well is harder than physics is because while the physicist can be corrected and inspired by experiment at every step of the way, the metaphysicist is on his own. One part of his conclusions can be matched against physical theory; his premises and intermediate steps he must make do as best he can; it is thus very likely that he will make mistakes, and those mistakes will be difficult to find.

The problem is, though, that although to do metaphysics well is very difficult, to do it badly is exceptionally easy. For this reason, there is a lot of bad metaphysics around, both in those who don't tend to think about things in philosophical detail, and even in some who turn their love of wisdom into a profession. If you think about questions of God, or of ethics, or even politics, but are not consciously thinking about metaphysics as well, you thoughts will be based on metaphysical assumptions, but because you don't think about it they will almost certainly be incorrect assumptions, which means that you will get everything else wrong as well.

Can we ever be certain of what the true metaphysical theory is? And if not, is that an argument against spending time on it? I find it useful to outline five perspectives through which a metaphysical question can be viewed (we may, of course, be in different categories with regards to different questions; and I make no claims that this list is exhaustive).

- The decided dogmatist states that this is correct; we know the truth and that is it. No more questions.
- The undecided dogmatist has reduced the possibilities to a small number of (possibly related) answers, but believes that there is currently insufficient evidence to choose between them.
- The honourable agnostic says we do not know the full truth, but we are going to strive to find it.
- The dishonourable agnostic says that we do not know the full truth, but we are satisfied with ignorance.
- The inconsistent agnostic says that we cannot know the truth in any respect (or do not need to know it); there is no point in studying metaphysics.

I distinguish the last perspective from the second because the last is more concerned with our limitations and states that we can know nothing; while the second states that we know a great deal, and can in principle know more, but we do not yet know quite enough due to a lack of evidence but not any lack in the ability of reasoned thought or our own capability to sustain it. I label the last viewpoint as inconsistent, because 'We cannot know the truth about metaphysical statements' is itself a metaphysical statement which is claimed to be true; it divides possible answers into two categories; those we can understand and those we

cannot; and states that the true answer is in the second category and not in the first. Thus he has made a profound proclamation on the basis of no inquiry or evidence whatsoever. A similar argument applies to he who says he has no need to know; he needs to know it (at least in sufficient detail) in order to know that it is not worth knowing. To my mind, we should aim for the first perspective, and how blessed are those who, having completed the labour, have reached it! (as long as they got it right); while the second and third are points on the way towards it (and a perfectly reasonable place to be, and the place where most of us, including myself, actually are). As a lover of wisdom, and one who believes that we should use our intellects to their fullest ability, I find the fourth perspective to be deeply unsatisfying, while the last is deplorable. At least with regards to questions of religion, the atheist is a dogmatist, and the Christian is a dogmatist; but both almost invariably the second type of dogmatist. Few people claim to know the full truth, and I doubt that anyone who makes that claim today is correct to do so. But we can know the truth in part, within certain bounds, and that is frequently enough to draw useful conclusions.

So why worry about metaphysics? Because it is impossible to do without it. Almost everything we believe or trust in ultimately boils down to a set of assumptions which are drawn from a metaphysical world view. We might not realise that we are doing that; we might not be aware of our assumptions or believe them to be self-evident, but ultimately they are always there and unless we examine and test them, that is to say study metaphysics, we ought to have no confidence that we are right about anything, or at least anything of importance. But surely physics, derived from experiment (along with chemistry, biology, etc.), can be done without metaphysics? No, because the scientific method requires certain assumptions, such as the validity of our senses, the existence of an external reality, the applicability of mathematics to a description of nature and the objectivity of our experiments, and these assumptions are ultimately metaphysical.

Can we make do with the minimal set of assumptions required for the empirical method, and be satisfied with that? No, because unless we investigate those assumptions, and what is connected to them, and what denied by them, we will not know how to interpret our experiments: are our conclusions truly drawn from experiment, or are they drawn from the presumptions that we put into the experiment? Are those presumptions truly necessary for the experimental method? Could another set of assumptions do as well? What other conclusions do the presumptions imply? Are the presumptions valid? All these questions – necessary to answer before we conduct our first experiment, if we want to be rigorous – are metaphysical investigations. We cannot sidestep the metaphysics by the experimental method without begging a lot of questions, and questions begged are often questions answered incorrectly. And, of course, metaphysics is not only the basis of physics, but also the other practical branches of philosophy. We need to try to connect these to ensure that our entire world-view is consistent.

But does not the success of science demonstrate that much progress can be made without metaphysics? After all, most scientists know next to nothing about the fundamental philosophy, and even those who do disagree among themselves. But these scientists have inherited a methodological tradition which dates back to Medieval Europe, and the founders of that tradition did study and understand metaphysics. It is possible to continue to work taking as premises the conclusions of those early pioneers without worrying about how those conclusions were drawn. Possible, but worrying, because eventually one needs to check that the conclusions are consistent with those first premises. We need physicists who are philosophically aware, or philosophers who are aware of physics and communicate with physicists, otherwise the whole adventure is in danger of getting off-track by losing the link between first premise and conclusion.

For example, to address the question of God requires metaphysics; 'God exists' and 'God does not exist' are both metaphysical statements. If we claim that we can address this question without metaphysics, we will still be doing metaphysics, but almost certainly

doing it badly because we are not thinking about it rigorously. We must therefore start by constructing the necessary definitions required to frame our discussion, and this is what I do when I turn to the question of God. Problems in ethics also require that we have competently developed our metaphysics: 'What is good?', which is the starting point of ethics, is another metaphysical question.

I will conclude this introduction by saying that not everyone will agree with my statement that the metaphysicist has no outside knowledge to guide him, analogous to the physicist's experiment; in fact I disagree with that statement in particular circumstances. If we consider metaphysics as solely a secular discipline I believe the statement to be correct; as a theist I believe that no discipline should be regarded as entirely secular, and this least of all. I believe that the discussion will lead us to something we call 'God', and that God is capable of and has taught us about Himself through revelations at various times in human history[1]. This offers us with a corrective, as the experiment corrects the physicist. We do not need this revelation to make a great deal of progress in metaphysics, and I will not use it here. But to reach the full truth about God direct information from revelation is necessary. I do not believe that we need revelation to prove the existence of God, or at least demonstrate it to be a highly plausible possibility. I do believe that we need it to find the Christian God (or indeed the Islamic god, or a host of other possibilities). Reason by itself can tell us where to look, and how to interpret what we see, but only by looking and seeing can we find the answer. I (or rather, those I am following) used revelation as a scaffolding to support the structure as it was being erected; once it is completed, the scaffolding can be removed, and the building stands on its own.

3.2. Assumptions and definitions

DEFINITION 3.2.1. **Rationality**:
By rational, I mean that object under study obeys the laws of logic, and most particularly self-consistency. △

Discussion:
The rationality of nature and supernature is a necessary assumption if we are to apply our reason to understanding them. It cannot be proved (Gödel's theorem), although if it were not true one would expect the consequences to be obviously apparent, since a break down in logic in one place would lead to inconsistencies elsewhere (errors multiply). I will assume that it is correct without any further discussion. □

DEFINITION 3.2.2. **State**:
A state describes the particular circumstances in which a being could happen to find itself in in any particular moment. This might be described by a list of the properties of a being when it is in that state, though I prefer to keep the states more fundamental and derive properties from knowledge of a state. In a mechanical picture, a state would be a complete description of the structure, internal motions and matter of the object at a given moment (and each state uniquely described by a single location, momentum, and any other relevant qualities such as the type of particle); in an Aristotelian system it corresponds to one of the potentia, and provides a description of the abstract part of the being in one particular configuration; in a quantum system it is represented by a single eigenvector of the effective

[1] *Revelation* literally means an unveiling, or something that provides evidence for truths not previously accessible to us. I shall use the word to denote special or positive revelation, i.e. arising from particular acts of God that go beyond the usual concourse of nature.

Hamiltonian describing the being (that will make more sense later). Every actual object must be in one state or another and (although I reserve the right to question this part of the definition later) cannot be in two different states at once. For example, a man might be standing, sitting, or lying down. Sitting is one state of the man; standing another, and lying down another. The man cannot be both standing and sitting at the same time, but he must be in either one of these or some other posture. Δ

DEFINITION 3.2.3. **Principle of sufficient reason**:
For every being whose present state is not logically necessary given the being's existence, i.e. it could potentially be in more than one particular state (including non-existence), each of which is logically consistent, but in actuality exists in only one of those states at a given time, then there is an explanation why it is in that particular state rather than another, which we can in principle (though not necessarily in practice) discover. Δ

Discussion:
This principle is controversial, and many philosophers reject it[2]. It is necessary if we are to come to a rational explanation of the universe; if things happen for no reason – which is the same as entering a physical or metaphysical state for no reason then obviously our ability to understand nature will be limited up to that point where there is no rational explanation. To deny the principle of sufficient reason is thus in effect to deny the rationality of the universe, or at least to suggest that it lies beyond the capability of not only our intellect but any possibility of intellectual inquiry. It is, in short, to begin a quest building in the requirement that it will ultimately fail. I cannot see how it is possible to have a rational universe without adopting this principle; indeed I believe that physics contains it as a paradigmatic assumption, since otherwise we could just stop somewhere and say 'This is all we can even in principle know,' regardless of whether or not that statement is true.

I have left a little bit ambiguous what I mean by an 'explanation', and I want to give an example to clarify one aspect of what I mean. Suppose, for example, we have a radioactive atom which emits a beta particle (an electron), and turns into a different sort of atom (ultimately from a neutron decaying into a proton and electron, or, even more ultimately, a down quark decaying into an up quark and a W^- Boson which in turn decays into an electron and neutrino). It is often said that this event occurs for no reason: there is no particular reason why it happened at one particular time rather than another. Rather, there is a certain probability that a down quark in a particular type of atom will decay into an up quark, electron and neutrino in a particular span of time, but we cannot say, even knowing the full state of the universe, whether it will decay then or it won't. But the principle of sufficient reason, as I have phrased it here, does not say that there is an explanation for every *event*, but only that there are explanations for material *states*, and, if there are immaterial beings, perhaps their states as well. It is saying that there is a reason why there is an electron (and neutrino) with a particular momentum and spin; the reason is that it came from the decay of a W^- particle with a particular momentum and spin into the electron and neutrino. Why was there a W^- Boson in that particular state? It came from the decay of the down quark into an up quark and the W^- Boson. And, of course, we can continue in this way, explaining the origin of one material state in terms of a different material state, back to the beginning of the universe. The second part of the explanation are the laws of physics which underlie this whole process. The principle of sufficient reason states that, if the principle is true, we will never be able to say of any particular electron, 'Well, it just appeared out of nothing.' (By nothing I do mean nothing, and not a quantum

[2]See (Pruss, 2010) for a recent defence of this principle, and response to various objections levelled against it.

vacuum, nor the laws of gravity, nor whatever is proposed in the next popular work that attempts to explain how something came from nothing by redefining the meaning of the word 'nothing.')

With regards to physical objects (and here I will assume that we can neglect any supernatural intervention in the universe), the idea of conservation laws implies that the principle of sufficient reason is correct. For example, we have the conservation of momentum (when I refer to the conservation of momentum, I will always mean the conservation of four momentum, i.e. both the conservation of energy and momentum as described in Newton's physics), the conservation of total electric charge, and so on. When we look at a particle and ask 'Why does that have the momentum it does?' the answer is never[3] 'It came from nowhere', but instead 'The momentum came from some other particle which previously carried that momentum.' That implies the principle of sufficient reason. I know of no physical theory consistent with experiment that violates the principle of conservation of momentum; and for that reason with regards to physics I am confident that the principle of sufficient reason holds.

More controversially, we can ask 'Why is the universe itself in the state that it is in?' or 'Why are the laws of physics as they are?' In our current understanding, the laws of physics are described in terms of a quantum field theory constrained by certain symmetries, choices of fundamental parameters and a few other conditions. The question concerning the laws of physics then becomes 'Why a quantum field theory?' 'Why those symmetries?' 'Why do the parameters take the values that they do?' It is in answering these sort of questions that the principle of sufficient reason becomes controversial. These questions go beyond physics into metaphysics. Physics itself cannot answer them (physics instead poses the questions in their most accurate form); that does not mean that they cannot be answered. Every answer to these questions, even 'There is no answer' or 'It just happens to be that way' implies a metaphysical framework; and that framework may either be consistent with the principle of sufficient reason or inconsistent with it. I will assume that the true metaphysical framework is consistent with the principle of sufficient reason. One reason for this is the obvious statement that our metaphysics must be consistent with our physics (in principle we ought to be able to derive at least some aspects of the laws of physics from our metaphysics, and in this way we can test between different metaphysical models). It seems to me highly unlikely that a metaphysics which is ultimately in violation of this principle can lead to a physics which satisfies it: usually an irrational theory does not lead to a rational answer, but only more irrationality. If one makes a mistake in one's algebra and introduces a contradiction, that contradiction remains throughout the calculation, and only multiplies rather than reduces (the only exception I can think of is where we can get a rational answer from a irrational premise, without introducing a second mistake which by the most unlikely of chance happens to precisely cancel out the first, is when our calculation is ultimately of no consequence. For example, when we assert $7 = 5$ and then multiply both sides of that equation by zero). And, of course, if any of these questions do turn out to be unanswerable, it doesn't violate my definition of the principle. 'Why a quantum field theory?' is not a question concerning a particular material state. The universe is not an individual being.

The principle of sufficient reason thus states that the sequence of explanations has a natural termination, and that this terminus is not something that could in principle be in a state other than what it is. If the principle holds, then it is in principle possible (though obviously it might be beyond our own limited reasoning abilities) to by reason reach a foundation of everything else in the universe, if the assumption concerning the rationality of nature and supernature is valid. If the principle does not hold, then either there must be a point where the chain of reasoning will fail, and we will have to say that the universe happens to be in this state for no reason, or there is no termination to the chain of explanation.

---

[3]The word *never* might be a bit controversial here, e.g. what about the Big Bang? What about particle/anti-particle pairs emerging from the vacuum? I'll turn to these questions later.

Numerous arguments have been given in support of the principle of sufficient reason. The principle denies that there is inexplicability. So, for example, given two substances A and B, identical in every detail relevant to whatever measurement we make, put in the same circumstances, we would expect that they would have the same probability of having the same outcome, and those probabilities for the outcomes would be different if we started with something other than our substances A and B. This is because there is no reason why they should differ; such a difference in outcome from identical circumstances would be inexplicable. But this expectation relies on the idea that there are reasons for the effects which can be traced back to the original causes. Indeed, the scientific method assumes the reproducibility of experiments; were two identical circumstances to have different probabilities for the same outcome then this foundation of the scientific method would fail. But this is to say that in this circumstance, this lack of inexplicability implies that there is a reason for the various outcomes of the experiment to have certain probabilities; with the reason contained in the original set-up. If the principle of sufficient reason is violated, an effect might be traceable back to A, but there is also the possibility that something else could happen for no reason. And since there is no reason, we can't say that the same irrational effects would arise from the experiment that started with B. If the effects can't be traced back to A and B, but they are nonetheless the same, then that commonality implies that there is a reason why the effects are the same in the two circumstances, not A and B but something else that connects the two experiments, which we can then investigate and discover. In other words, a violation of the principle of sufficient reason would imply that when we have identical circumstances we ought to expect a different probability of a particular outcome; and this goes against all our experience and offends our rational sense.

To deny the principle of sufficient reason is to say that there are some things which can be explained rationally and others that cannot. It is obvious that there are things which admit a rational observation; every observed physical process consistent with theory is an example. Suppose we assert that there is also something inexplicable. This is not a statement that can be proved: the best we can say is that we do not currently know the reason, which is very different from saying that there is no reason in principle. We all know that, as our knowledge has increased, things which were inexplicable according to past knowledge are now understood; it would be foolish to say that those things which are badly understood now can never be known in the future. That means that the denier of sufficient reason must draw a line somewhere to divide those things which are rational and those which lack any rational explanation. But where should we draw that line? How do we know we have drawn it in the right place? What is different between those things which are not rationally explicable compared with those things which are? How do we know there is such a line? The simplest solution is to say that there is no such line, and those who deny the principle of sufficient reason need to provide a convincing explanation of why there is such a metaphysical boundary; why some quarks or gravitational singularities plopped out of nothing, while every existent quark and singularity we observe has an efficient cause.

Let us suppose that there is a terminus of explanation which is not an ultimate principle which logically cannot be other than what it is. Suppose that there is a being X which exists as the terminus of the chain of reasonable explanation which could be in various different states X_1, X_2, X_3 and so on. The chain of explanation terminates with it in state X_1. We can also imagine a similar being which decayed into the state X_1 from the state X_2. These two hypothetical beings in state X_1 are identical in every way, so it is unreasonable to say that one of them can serve as a brute terminus of explanation, and the other one can't. The only possible terminus of the chain is something that can't in principle be caused, that is to say something that can't in principle change into a different state or emerge from a different state. One cannot just draw the line where the terminus of explanation is reached in an arbitrary way without making those cases where we can rightly use the argument also

invalid. Why should we say that X_1 and X_1 are different for no reason and still draw the inference (which we all do) that A and A are not different for no reason?

The other option is that the chain of explanation extends back to infinity. But this would be a countable infinity, which creates numerous mathematical paradoxes (Spitzer, 2010). Thus we can also rule this option out. If so, then it the principle of sufficient reason is correct, and we can in principle (though not necessary in practice) come to a final understanding of the universe that is not just a shrug and 'Well, that's just the way it is.' □

DEFINITION 3.2.4. **Brute fact:**
A brute fact is some piece of data which could in principle take a different value but there is no underlying reason that it takes the value that it does. △

Discussion:
An alternative definition of the principal of sufficient reason is that there are no brute facts.

Of course, there is plenty of data which we cannot in practice explain; there might be data where we will never, no matter how much effort is put into it, be able to find an explanation, due to our inherent limitations. However, a brute fact is a stronger condition than either of these: it does not depend on ourselves and our limitations, but on nature itself. It is something which is in principle unexplainable; it is logically impossible that an explanation could be found. Also note the importance in this definition of saying that something could be different than it is; an object which can only be in one state needs no explanation of why it is in that state and not another, so is not a brute fact. □

DEFINITION 3.2.5. **Realism:**
I define realism as the belief that the world is objective, independent of the human mind, and contains universals (that is certain concepts in nature which apply to numerous different objects; for example biological realism states that there is a universal *cat* which describes commonalities held by all cats, and there is a universal *unicorn* which can be applied to all unicorns – although it happens that there are no unicorns in practice, a unicorn remains a logical possibility, and so we can identify logically possible specific instances of unicorns with the universal unicorn which describes the species as a whole). The principal alternatives to realism are nominalism, which holds that there are no universals in principle; conceptualism which believes that universals only exist in the human mind, as a useful bookkeeping tool; and idealism, which holds that there is no objective existence at all outside our experiences.

Traditionally, the main forms of realism are monism (the belief that there are no distinctions at all; everything is an universal), extreme realism (Plato's theory), which states that universals have an independent existence, and moderate realism (derived from Aristotle and perfected in the high middle ages) which states that universals exist not in themselves but only as part of a substances. So, for example, a crystal of copper is the union of the matter (which distinguishes one lump of copper from another) and the form (which each lump of copper shares; a description of the form includes a description of the possible arrangements of the various fundamental particles to create first copper atoms, and subsequently how the atoms are arranged to form the metal). Properties are subsequently derived from the form; so the redness of two objects are ultimately derived from the two objects having similar structures of the electron energy bands, and thus the two objects can be said to both display one and the same colour rather than two colours which happen to be similar. In this work, I will advocate a form of moderate realism, and when I refer to realism I shall generally mean this type of moderate realism.

I am also a scientific realist, in that I believe that the objects described by the mathematics of physics have objective reality and are not merely a bookkeeping trick. For example, in the nineteenth century, it was debated whether the then unobservable atoms had any objective existence beyond their place in the equations; now the debate moves down levels to things such as quarks and virtual particles; a scientific realist would say that the equations describe objects existing in reality; the intrumentalist that they are just useful fictions that happen to make the equations work but don't exist in practice. The main argument for scientific realism is to ask why the equations work if their subjects are merely fictitious devices rather than genuine components of directly observable matter. Δ

Discussion:
I believe that realism, while impossible to prove against idealism, is required for physics to have any meaningful interpretation (equally, idealism, the denial of an objective world, is impossible to prove). At its core, physics describes how material objects interact with each other and expresses these interactions in a mathematical structure. This physical law can then be used to predict the behaviour of objects in similar circumstances. Physics involves the categorisation and application of these laws. These laws govern all material objects, regardless of whether they are being observed, and applies equally to all objects of the same type. Thus one electron will behave identically to another electron; each pion has an identical mass and electric charge and all pions are interchangeable, and this is not merely a matter of human categorisation but the nature of the particles. This law is inherent in the nature of the particles; for example if you place any electron in an electric field it will behave in a certain way, while any neutron would behave in a different way. That these laws apply equally to every instance of one type of matter thus rules out nominalism; while these laws are assumed to exist in the same way before there were any human observers, ruling out idealism; and the mathematical formalism describes real aspects of nature that exist independently of our formalism (the interaction of particles is identical despite how we express it — Feynman said that every good physicist knows at least six or seven different ways to describe the same physics (Feynman, 1965): the mind dependent formalism differs; the behaviour of the electron does not), in contradiction to conceptualism. Idealism should be rejected, because some explanation needs to be found for why there seems to be a real world out there, including consistencies which only reveal themselves to our understanding after the fact. The most obvious answer is that the appearance of something which cannot be distinguished from a physical world is, in fact, caused by there being a physical world out there: any other explanation has far less explanatory power.

Before I move on to my example, I need to give a brief introduction to the atom and the quantum particles that make up the atom. I say 'quantum' particles to distinguish them from classical particles. Classical particles are almost certainly what you think of when you hear the word 'particle': individual localised lumps of matter. The word 'quantum' in front reminds us that at the sub-atomic level, things follow very different rules and our classical intuition about particles is not correct. A quantum particle is different, though I will defer my discussion of how it is different until my full discussion of quantum physics.

As is well known, material substances are composed of atoms, and those atoms are divided between a nucleus at the centre, which is very heavy, and a cloud of very light electrons surrounding it. Think for a moment of the solar system picture, with a very heavy star at the centre and light planets orbiting it, and it will give you a rough idea of the atom. Then forget about the solar system, because once you have that idea of a stationary nucleus at the centre and electrons moving around it, it is more misleading than helpful. The solar system is governed by the laws of classical mechanics (specifically gravity) and the atom by the laws of quantum mechanics (quantum electrodynamics), and the two are very different from each other. The electrons do not move in fixed elliptical orbits, and the reality is impossible for us to imagine, though not impossible for us to describe and understand

with our intellect: to understand we need to move from mental pictures (which in quantum physics get in our way more than they help us) to the more abstract language of logic and mathematics. A quantum particle in a defined state is not the single point-like classical particle, but rather it has a non-localised amplitude, spread out over an area, but with a well defined spatial distribution. Different distributions correspond to different energy levels. Now the nucleus is composed of two types of quantum particles, the proton and neutron, with the atomic number counting the number of protons in the nucleus and the atomic mass (to a good approximation) the total number of protons and neutrons. Protons and neutrons themselves are not fundamental particles, but are composed of three particles called quarks. The interactions between these fundamental particles are conveyed through the creation and destruction of force carriers, photons which carry the electromagnetic force, the W^+, W^- and Z^0 Bosons which carry the weak nuclear force, and the gluons which carry the strong interaction. The electromagnetic force is primarily responsible for the interactions between the nucleus and the surrounding electrons, and ultimately for chemistry, electricity, magnetism, and most things we see on a large scale (the only exceptions are the effects of gravity). The strong nuclear force binds the protons and neutrons together in the nucleus. The weak force converts one type of quark to another, and is responsible for certain types of radioactive decay. The quarks also combine to make other particles apart from protons and neutrons, the pion, rho meson, and delta particle, for example, but these are less significant for our every-day lives. We are almost entirely made up of protons, neutrons, electrons, and the force carriers.

Now for my example: let us consider a simple experiment. We have two beams of quantum particles (for example, electrons, protons, alpha particles, or something of a similar scale), one coming from point A and another from point B some distance from it. They travel towards each other with equal speed and beam intensity but opposite velocity, and therefore 'hit' each other at the mid-point of the line AB, a point which we call C. We then also place two detectors at points D and E, where the midpoint of DE is also at point C, and there is a $90°$ angle between AB and DE. This sort of experiment is routinely carried out in particle accelerators. We can, in principle, using the mathematics of quantum mechanics, calculate the probability that particles will hit either of the detectors. So, for example, if we fire an alpha-particle in one beam against a proton in the other, then we will detect a certain probability that the alpha particle would hit one of the detectors, and a certain other probability that the proton will hit the detector. Combining this probability with the intensity of the original beam, we can calculate to a good precision how many particles will hit the detector in a given time; and, of course, the prediction from the theory and the results from the experiment agree to within the limits of experimental, theoretical and statistical uncertainty. In practice, the theoretical result will be described in terms of a probability distribution (this would be true even in classical mechanics, since there would be uncertainty over the precise location of each particle within the beam and thus over how the two beams interact with each other; the uncertainty in quantum physics is intrinsic to the theory and can't be avoided even if perfect knowledge of the initial conditions were possible). So, in a certain time interval, we can calculate that there is a certain probability that we will record no particles at the detector, a different probability that there will be one particle recorded, a further probability that there will be two particles, and so on. We can calculate these probabilities from the theory, and then perform the experiment for enough time periods to see how well the prediction matches up to reality and I know of no experiment of this type where the most modern theory has got it wrong.

Now, the mathematics of quantum mechanics depends on whether the particles are identical or not. For example, a certain class of particles, called Fermions, obeys what is called the Pauli Exclusion Principle, which says that two particles of the same type with the same quantum numbers cannot be in the same state at the same time. This exclusion principle is of huge importance. Without it, there would be no chemistry. Electrons are

a type of Fermion, and their states are distinguished by energy, momentum, angular momentum, and something rather unfortunately named spin, which has two possible values called spin up and spin down. There are certain allowed energy levels around each atom. Particles naturally head towards the lowest energy level, in this case known as the $1s$ shell. If there were no exclusion principle, then every electron will fall into the $1s$ shell; but as it is, there can only be two electrons in the shell, one spin up electron and one spin down electron. Additional electrons cannot fall into the $1s$ shell, but must gradually occupy the higher energy shells, which are labelled by atomic physicists and chemists $2s$, $2p$ and so on. Thus, the distribution of the electrons around the atom depends on the number of electrons, which is the atomic number, and this is why different elements have their various chemical properties. Without the exclusion principle, there would only be $1s$ electrons, and only interactions between atoms involving $1s$ electrons would be possible – very different from the real world. But again, I emphasise that the exclusion principle only applies if the particles are indistinguishable from each other – if they are of the same physical type. If they are in any way different, then the exclusion principle does not apply, and there is nothing in quantum mechanics preventing them from all dropping down to the $1s$ level. There is nothing in quantum mechanics stopping the lowest $1s$ energy level of Beryllium (atomic number 4) containing two electrons and two muons (a particle with similar properties to the electron) – at least until the muon decays. The exclusion principle only applies to indistinguishable particles. There is also a second class of particles, known as Bosons, which do not satisfy the exclusion principle, so you can pile as many of them as you like into a single state.

Now, the exclusion principle also applies in the experiment described above. If you used two spin up electrons in the experiment, then the result is dramatically different because the two electrons cannot be distinguished from each other. We get one probability distribution if the particles are distinguishable, and another distribution if they are indistinguishable (in fact, if the particles are indistinguishable Fermions, the probability that they scatter at 90° is zero; for indistinguishable Bosons the expected number of particles is twice what you would obtain if the particles were distinguishable). This difference can be predicted in the theory and then measured in an experiment – and the theory gets it right. Because of conservation of momentum, if one of the particles hits detector D then the other one must be recorded at detector E. Because the particles cannot be distinguished, we cannot tell whether it is the particle originating from A which hits detector D or the particle originating from B, and this makes a difference to the mathematics of quantum theory. For example, if we in some way 'painted' an α on one particle and a β on the other particle so that we could tell them apart, but made no difference to the mass, charge, spin or anything else that affects the forces between the particles, then the results of the experiment would be different than if we did not do so.

Suppose that we performed the experiment where the detectors cannot distinguish between the particles; they seem to be indistinguishable, and we get the corresponding intensity at each detector. Then, suppose that we replace the detectors with some that can tell the difference between the particles. We would now, by the mathematics of quantum mechanics, get a different intensity of particles at each detector. But this is impossible: the detector we use cannot affect the effects of the collision occurring at C. This means that it is not possible, even in principle, to tell which of the spin up electrons at the detectors came from A and which came from B. The indistinguishably of the particles is built into the laws of physics; it is not a matter of our knowledge being imperfect; it is not a feature of our experimental set-up, because the laws apply both inside and outside the laboratory: the same behaviour occurs whether or not we are conducting an experiment. The mathematics only works if we treat the electrons (or Bosons) as two separate instances of the same type of particle, and fails if we treat the electrons as two separate individual types of particle which happen to have similar properties.

Of course, different types of particle also have different masses, electric charges, and so on. Is this enough to explain the experimental result? Is it merely the difference in mass which causes the different experimental result? A different mass certainly affects the probabilities, but not in this way. Such qualities and differences are accounted for in the theory, but without the exclusion principle they are insufficient to lead to the observed effect. It is not the identity of properties that causes the observation, but that the particles are the same type.

Thus if nature itself can distinguish between different types of particles, and this has experimental consequences, then there must be different types of particles in reality, and not merely as our own bookkeeping measure. Translating back into the language of the philosophers, we can say that these types represent universals. Thus there is a universal representing electrons, a universal representing protons and so on. Thus the existence of at least some kinds of universals is shown from observation, demonstrating philosophical realism, at least at the level of fundamental physics.

Compare this to what we expect from nominalism or conceptualism. Here, at best we class the particles into the same type just as a matter of convenience because they just happen to have similar properties. But in reality, there is no universal 'electronness' describing the two particles. In nature, the two particles are distinct, even if we ourselves cannot distinguish between them. Under nominalism or conceptualism, we would not expect to see the Pauli Exclusion Principle, since nature would not be able to distinguish that two electrons were both electrons and not two separate particles which happened to share similar properties. We would expect the same distribution at the detectors whether or not the particles appear to be the same or not. But this is not what we observe. Thus the similarity between the particles is not just a matter of our convenience, but truly built into nature – in agreement with realism and in contrast to nominalism and conceptualism.

Nominalism and conceptualism gained popularity when it seemed that classical mechanics described fundamental physics, and in terms of classical mechanics there is no strong physical argument to oppose them (although there are various philosophical objections). In classical mechanics, the only things that matter are the properties of the particle, such as its mass and electric charge. There is no exclusion principle in classical mechanics. Two similar particles behave in broadly the same way regardless of whether or not they happen to be of the same type – so whether or not they share a universal form makes no difference to the predictions of classical physics. It is therefore quite plausible that different classical particles are distinct and only happen to share similar properties, and so are grouped together only for our convenience. Up to the early twentieth century, these world-views were not unreasonable, and not surprisingly many philosophers adopted them.

However, we now know that the Newtonian picture of the world is wrong. I do not believe that nominalism or conceptualism are tenable from the perspective of quantum mechanics or quantum field theory. They should have been abandoned (but have not been, perhaps due to the unfortunate divorce between metaphysics and physics that occurred around the time of the emergence of quantum theory). We have to admit to the existence of universals, at least at the level of atomic and subatomic physics; and since the properties of materials (such as colour, rigidity and so on) are derived from the energy levels of the materials, which are computed ultimately from the interactions between the fundamental particles, we have to adopt a realist understanding these properties as well. The colour of one lump of plastic is the same as the colour of anything else with an identical structure of energy bands. The identical natures of these objects are not just accidental by-products of our classification procedures, but inherent in nature itself.

This will be a common theme throughout this work: what was plausible under the world-view of Newtonian mechanics has now become implausible in the quantum world-view, but much of modern philosophy is based on Newtonian mechanics. Nominalism and conceptualism make sense if Newton's theory is accepted, and perhaps fit better with it than

realism. And it was on the basis of Newton's theory that the 'great' philosophers of the eighteenth and nineteenth centuries formulated the modern conception of these world-views, and on this same basis that they came to be viewed with favour. But Newton's theory is wrong - and it is wrong in a way that is entirely consistent with classical realism, but not with its more modern competitors.

The question, of course is different when we move beyond physics. If we are realists in physics, then we also have to be so in chemistry, since if the building blocks of chemistry are indistinguishable, then so are what is made from them. One ethanol molecule is identical in substance to any other ethanol molecule[4]. However, it is not so clear when we try to apply this idea to biology in the customary way. For example, we classify tigers as a species, suggesting (to the realist) that each individual tiger shares some features that together we may call 'tigerness', while everything that isn't a tiger lacks 'tigerness'. Clearly, the above argument does not apply in this case, since 'tigerness' is used to classify similarity rather than indistinguishability. Therefore, at the level of biological species, the argument I have used to defend realism in physics and chemistry doesn't apply. One can certainly say without contradiction that we can correctly identify universals in physics and chemistry, but not in other fields of study. However, it is not obvious that this view is correct: we can also honestly say that there is a universal called 'tigerness' which we can discern through observation but also exists apart from our observation. Even in this case, the realist position seems to be to be considerably stronger. We know, for example, that each tiger has a certain genetic similarity to every other tiger. In the space of all possible DNA codes (by possible, I mean in terms of natural selection, so well adapted to survival of the species when the environment is in the an equilibrium state), there is a small region we can associate with the modern tiger, and this region is surrounded by a space which is not presently occupied by any living organism (so animals with that DNA would not be sufficiently well adapted to pass on their genes when in competition with the modern tiger). This region of DNA is objective; it is an abstract concept existing even apart from our observations. We can therefore point to some continuous entity in nature, and say that it corresponds to 'tigerness'. But this is just what we mean by a universal. 'Tigerness' is not merely something we invent, but a concept we use to describe something that exists objectively in nature. Therefore, even here, I think that realism is a far stronger position to hold than nominalism or conceptualism[5].

The question of universals is how can we explain that certain objects both seem to be of the same type, and yet are also distinct. This involves two concepts. Firstly, that there is a universal or unifying principle connecting every member of the same type. Secondly, there is a distinguishing principle marking them out as individuals. If nominalism, the denial that there is a unifying principle in matter, is untenable, what about its opposite, monism, namely that there is no distinguishing principle, and that consequently everything is ultimately the same object? But this runs against all our observations. I am distinct from my brother and neighbour; I have different thoughts, different actions, a different location, and am made of different ratios of the atomic elements. So are there any possible middle grounds? Plato suggested that the differences were caused by a negation of being; objects were different because certain aspects of a beings true nature are missing or distorted. Both I and my neighbour fall short of the ideal man (the universal), but we fall short in different ways, and can thus be distinguished. However, this answer is also unappealing, because some of the differences between objects are contradictory; they cannot be due to an absence of qualities possessed by the ideal because the ideal can't simultaneously have both the features present in other men which I lack and the features which my neighbour lacks.

[4]I do think that the various oscillations in the positions of atoms are important; one cannot say they are identical in the sense that they are in the same state. However, they are identical in the sense that they have the same set of possible states. As shown below when I define it, the concept of substance can account for minor variations such as this. That they are of different states does not affect the statement that they are identical in substance, as will be clear when I define the concept of substance later.

[5]See (Oderberg, 2008) for a defence of realism with respect to biological species.

Aristotle answered the question by supposing that objects are composite, the fusion of two parts: one part corresponds to the unifying principle, e.g. something which all men have in common, and the other to the distinguishing principle, what separates one man from another. Thus things can be both the same and different simultaneously, because they are the same in one respect and different in another respect. The unifying principle he called *form*, and the distinguishing principle *matter*. Having eliminated the alternatives, we are forced to accept Aristotle's solution, or at least something built along similar lines.

There are, of course, numerous different forms of realism. One question is whether the laws of physics are inherent within the material objects, or in some way external to them. For example, does the J/ψ meson (another of the zoo of subatomic particles) decay as it does because it is in the nature of J/ψ mesons to decay in that way or something external to the J/ψ acts on it to make it decay? Are the laws of physics a property of different types of matter, or something external to matter which it must obey? It is difficult to answer these questions just from a knowledge of the laws of physics. This is a matter of metaphysics, and there may be several metaphysical models consistent with physics. Nonetheless, we only need to consider realistic metaphysics.

I also reject empiricism, the belief that the only things we can know are what we observe. This is a stronger statement than to say that observations are essential for us to come to our knowledge; empiricism is a philosophy that denies that we have the ability to correctly reason from what we directly observe to what we have not observed or even cannot in principle observe. The one who accepts both reasoned argument and observation is not an empiricist. In empiricism, the laws of physics (since they cannot be directly observed, but only indirectly deduced from either observations or a priori principles such as symmetries) only describe and systematise our observations: we note certain regularities and call them a 'law', but in reality the law is simply a construct of our own, invented for our convenience but with no basis in reality. Instead I assert that they have an existence of their own outside of our categorisation of them. My reason for believing this is that empiricism does not explain why the laws of physics are obeyed, or allow for prediction. If the laws of physics are only a description of observed regularities, then they have no obvious validity when extended to things which have not been observed and are of a wholly different nature to what has been seen so far.

For example, we have discovered planets by noting the effects of their gravitational field on other planets. We see variations of the orbit, see that they can be explained by another planet, and the natural implication is that there is another planet there. In an empirical system, it does not follow that there is another planet there; the laws simply describe what we observe, and if we can describe the movement of the planets we observe with a mathematical pattern then there need not be anything unobserved to cause any deviations in the movement of the planets; indeed the concept of deviation makes no sense, since the only ideal we can compare the motion of the planets with is the actual paths that they take. There is no reason to suppose that there is another planet there (and that the variation of the orbit is not caused, for example, by an incompleteness in Newton's law) until we see it through a telescope.

Similarly, the Higgs Boson is required for the consistency of the standard model; before it was discovered it was widely accepted to be there; but if empiricism is true there is no reason for us to say that it is there until it was observed; the universe may as well not contain the Higgs. The ability of physics to predict things before we observe them thus is evidence that its laws do indeed represent something that has its place in reality, and are not mere human constructs to describe regularities.

Empiricism also does not explain anything; we knock two billiard balls together, and observe that they move in a certain way. Newtonian physics asserts that they must move in that particular way due to the laws of conservation of momentum and energy, as well as the details of their contact. But these laws are not something seen by experiment; in

modern physics they are deduced as a consequence of various symmetries of the action, another unobservable quantity. Empiricism asserts that their movement is just a pattern we see, and they could in principle move in another way. This completely fails to answer the question of why the balls behave as we observe; the most obvious answer is that there is an objective law built (in some way) into the structure of nature which compels them to do so. Empiricism rules out this answer not by valid argument but by assertion, and offers nothing in its place. Science is impossible in an empiricist world view, and its success inexplicable.

However, the opposite of empiricism, rationalism, which in its extreme form is the belief that only reason should be trusted, is also to be rejected. We need experiment as a both as a corrective to reason and to provide us with the initial premises. Reason can only argue from premise to conclusion; it cannot by itself tell us what the correct premises are or whether the conclusion is correct.

The best approach is the middle ground between empiricism and rationalism, not preferring either experiment or reason. Observation has its place, and reason has its place, so let us use them both in those circumstances which are appropriate for each source of knowledge. Accept that we can understand the true natures of things through our reason, but need observation to help us to do so.

Scepticism, which in its extreme form is the belief that no knowledge can be adequately justified, is without merit, in part because we obviously do already know quite a lot about how the universe works. But the position is also self-refuting: can the knowledge that no knowledge can be justified itself be justified? There is obviously a great deal we do not know, but to say that we cannot know it goes further than any evidence allows; and even if there limits to what we can know, that does not show that there we cannot know many important and profound things. On the other hand, the assertion that all our beliefs are adequately justified is foolhardy and itself unproven. Once again, a middle way needs to be steered between scepticism and over-confidence: we must be aware of both that we do in practice make mistakes, but yet we have the capacity to get many things right. The traditional arguments for scepticism fail because they do not take into account the idea that we can accurately parametrise uncertainty using probability: they assume that belief is either both accurate and perfectly precise (knowledge) or wholly imprecise, while in reality knowledge only needs to be accurate with specified imprecision.

The main philosophical families which accept realism are those of Plato, Aristotle, and the 'scientific' realism which is built on the mechanical world view[6]. The form is that which is shared between each object of the same type. The three realist philosophies differ by how they regard the form of large-scale structures. Plato believed that form is real, physical and precise; Aristotelian realism takes a middle position which affirms form as an abstract concept which is imprecise in the sense that there can be different manifestations of the same form, and takes everything we observe as a union between the abstract and universal form and the particular and physical matter; the mechanical picture accepts universals only at the level of fundamental particles (which, in some versions, are geometrical objects), which move according to fixed laws, and believes that everything else can be described in terms of the motions of these particles. These three options — that form exists and is precise; that form exists and is abstract and allows for variations; and that either form does not exist or is irrelevant to physics — provide a complete set of the available philosophies: we must choose one of them (although, obviously, each family is broad and a union of various related philosophies: there is no single Platonic way of viewing the world, and so on). I do not believe that Plato's system is tenable. I will therefore concentrate my discussion in

[6]Also worth noting, along with naive realism, are more modern constructions such as critical realism (Bhaskar, 1975; McGrath, 2001), modal realism (Lewis, 1986), process philosophy, and representative realism (the idea, associated with Locke, that we can only grasp representations of the true reality); however these tend to have a different emphasis to the question I am asking here. They are as much concerned with our knowledge of being as they are with my main concern, the nature of being itself.

later chapters on the Aristotelian and mechanistic philosophies. □

DEFINITION 3.2.6. **Statement**:
A statement is an expression which relates at least two concepts; either by comparing two different nouns (for example 'A horse differs from a dog'), or by attributing properties or actions to a noun ('This ball is red', or 'This ball is moving at 500 meters per second', 'This ball is about to hit that man' or 'I wouldn't like to be in his position right now') or to a verb, adverb or adjective ('Velocity is defined as the rate of change of position with respect to time'). Statements are the basis of all knowledge. △

DEFINITION 3.2.7. **Truth**:
There are several different senses of the concept of truth, and here I will concentrate on four of them. All these senses of the word truth relate to statements; it makes no sense to say that anything other than a statement is true (in these senses of the word)[7]. The first two are related to logic, and similar fields such as mathematics. Firstly, we have axiomatic truth, something which cannot be proved within a system, but can be treated as self-evident. I treat, for example, the principle of non-contradiction as one of my axioms. Ideally we want as few of these unprovable statements as possible, but they cannot be avoided altogether. Secondly, we have logical conditional truth, which means that if some premises are correct, then the conclusion is also correct. Consider, for example, the statement: 'If it is true that Socrates is a man, and if it is true that all men are Persians, then it is true that Socrates is a Persian.' This statement is, from a logical point of view, correct. It is a logical conditional truth: the conclusion follows from the premises. Of course, one of the premises is wrong, so the conclusion is also wrong in a different sense of the word 'truth', but it is true that if the premises hold, then so does the conclusion. The third and fourth senses of truth come from comparison of the statement with objective reality. The third of these is a statement that is always true (given our universe), analogous to the axiom in logic. One example is 'An electron always has a mass greater than the rest mass of an electron neutrino.' This is not a logical truth; we can define the concepts of 'electron', 'mass', 'rest mass' and 'electron neutrino' (and I here assume the standard definitions, which I will discuss later), but ultimately to see whether those definitions are useful, and correspond to objects in reality, we need to use observation. Obviously the use of the word 'electron' is entirely arbitrary (we can use any word to describe the same concept), but the concept itself is a part of nature and independent of our language; so it is useful to identify one word with it. Once we have identified the electron with a particular natural particle, then its form and consequently its properties are determined and fixed. Thus any (correct) statement about electrons in general, or any other universal, is a truth in this third sense. The fourth sense of truth is something which is both conditional and observed: it is only contingently true, because it depends on the context. For example the statement 'It is raining today' may be true one day and false the next depending on the weather and the location. There is no reason given logic or the definitions for it to be raining or not raining that day. The statement that it is raining is thus contingent, depending on circumstances, but can be compared against reality and shown to be either true or false. This comparison is objective:

[7]There are other senses of the word 'truth' which I shall in general not use here; for example we might say that a particular person is *true*, meaning that he is reliable, or that a line is *true*, meaning that it is straight. This does not negate my statement, since the same word in English is used to denote several different concepts; it is the underlying concept I am interested in ahead of the word.

it does not matter who is observing, or even if nobody is observing, the statement remains either true or false.

We thus have four senses of the word truth: logical and necessary, logical and conditional, observational and necessary, and observational and conditional.

Note that I am making one key assumption here – an assumption which depends solely on ourselves and thus which we must endeavour to satisfy – that our definitions are both useful and consistent (by useful I mean that at least one object in reality corresponds to the definition). For example, consider the statement 'It is necessarily true that flowering roses are sweet-smelling,' something which we can test by considering the consequences of the set of genetic codes that correspond to what we mean by roses; if all plants in that region of the genetic landscape smell sweet, then the statement is correct. However, it is the object itself which is essentially sweet-smelling, and the word 'rose' is merely our linguistic representation of that object. So, if I were to define the word 'rose' to mean what is commonly referred to as 'a pile of manure', then the statement is false. So is the statement 'Flowering roses are sweet-smelling' contingent on my definition of the word rose – and if so, there are no necessary truths, since all definitions are contingent? The point is that the definition and the statement must be taken together. Rather than 'roses are sweet-smelling', we should say

(1) A rose is defined as a flowering plant with a DNA structure within a certain region of the genetic landscape (I will leave it to the botanists to clarify what this region should be).
(2) All objects which correspond to that definition of *rose* are, once they have flowered, sweet smelling.

Those two statements together (combined with the other required definitions, and once we have identified the genetic region correctly) form a necessary truth. The definition, thus, performs part of the true statement.

Suppose we have a statement which could be true in the third and fourth senses. This can be categorised in one of four ways:

- **Nonsense:** A nonsense statement is one that is not internally self-consistent (for example 'This man is a penguin,' or 'This statement is false.')
- **Irrelevant:** An irrelevant statement is one that is internally self-consistent but cannot be compared with reality.
- **False:** A false statement is one that is internally self-consistent but does not correspond to reality.
- **True:** A true statement is one that is internally self-consistent and does correspond to reality.

Thus the concepts of truth and falsity are not the only options; if something is not false, that does not necessarily make it true. The class of irrelevant statements includes those conditional observational provided without the contextual conditions (such as 'It is now time for bed'). In this sense of the word, truth can only be applied to statements. This definition, of course, presupposes that an objective real world exists out there, and I do not believe it possible to have a coherent definition of truth in a idealist, subjective, nominalist or even certain types of empiricist system.

I should make one important qualification with regards to the expression *corresponds to* used above. Logical or mathematical statements, such as $2+3=5$ ought to be expressed precisely, if at all possible (it is not always possible). This is not in general the case for observational or empirical statements. For discrete quantities, such as integer numbers, a true statement can be precise (for example 'There are seventeen cows in that field') or imprecise ('There are between ten and twenty cows in that field'). For continuous quantities, such as location or length, every true statement must in practice be imprecise. For example, we cannot correctly say that 'This man's foot is 30.48cm long', because it will not be exactly 30.48cm long (to an infinite number of decimal places – as a physicist I would say that this

is a sub-manifold with a measure zero). Rather, we should always say 'This man's foot is between 30.47 and 30.49cm long', which may well be accurate. We can improve the precision of the statement, but we can never make a perfectly precise statement. This idea of *observational* or *experimental error* is perhaps the most important concept in modern natural science. A natural scientist who underestimates his errors cannot be trusted (at least as a scientist); and one of the most important and time consuming tasks of the scientist is to accurately keep track of and present the magnitude of all the possible sources of imprecision.

It is important to note that the concept that measurements of continuous quantities cannot be specified with perfect precision also extends to universals. We cannot say that a bear is precisely like this; rather there is a range of possible states associated with the concept of being a bear. Thus when we say 'This is a bear' we mean that it is one member of the set of states that is described by the definition of *bear*. In this way, we can both include similarities between different individual bears, differences within the species of bear, and differences between bears (as a group) and other animals within the same framework. We can, of course, speak in ways that are more and less precise: the different species of bear are universals which are more precise than the universal of 'bear', in that the sets describing those universals have fewer members; the universal describing each individual bear still describes a set of possible states of being, but is more precise still than the species; while the universal describing animals is less precise than all of them. △

DEFINITION 3.2.8. **Accuracy**:
A statement is accurate if it is true, and inaccurate if it is false, nonsense or irrelevant. Accuracy is an absolute quality of a statement; one does not in principle need to know any other statement (baring the necessary definitions) to judge whether a statement is accurate. △

DEFINITION 3.2.9. **Precision**:
Precision describes the range of values allowed for within a statement. It is in general a tool of comparison between expressions rather than an absolute measure. Thus we say that a statement is more or less precise than another statement; we do not (unless what we are comparing it against is clear from the context) say that a statement is precise or imprecise absolutely. For example, we may say that 'I walked between 2.5 and 3.5 miles today' is a more precise statement that 'I walked between 2 and 3 miles today'. However, the statement 'I walked for between thirty and forty-five minutes' is neither more nor less precise than the statement 'There are between one and two kilograms of oranges in the basket', since the two statements are expressed in terms of different units (unless we have some objective means of comparing minutes and oranges). △

DEFINITION 3.2.10. **Objective categorisation**:
To categorise something is to place them into various sets which emphasise their likeness of kind; for example we group numerous animals together as *tigers* and others as *bears*, with the proviso that a member of the group tiger cannot be a member of the group bear, and so on. We then categorise tigers, bears and many other *species* as animals as distinct from (for example) *plants*. All tigers are animals; that is that they share all the qualities that every animal shares; not all animals are tigers in that they do not all share the qualities that tigers share and other animals do not. Similarly, we categorise one species of sub-atomic

particles as *electrons* and another as *up-quarks*. That we do this, and that it is necessary for scientific study (otherwise experiments could not be repeated in different places with different members of the species) is not, I think, controversial. What is controversial, however, is whether this categorisation is objective or subjective. Most branches of nominalism would assert that it is subjective; merely the human mind incorrectly attributing patterns in nature; most branches of realism would say that it is objective and that we recognise something that previously exists in nature. (Obviously, we could say that some categories are objective are others subjective; I am of the opinion that everything that can be the subject of scientific study must be objectively categorised).

Objective categorisation then, is the belief that every being by nature is divided into certain kinds where there is a unity between members of each kind and disunity between members of different kinds; that these kinds may be further divided into species and subspecies (until eventually we reach sets just containing a single individual); and that these divisions are inherent in nature and physical law, and in no way dependent on ourselves[8]. △

Discussion:
Note that this does not mean that the categories we use are necessarily the objective categories that exist in nature: we may be in error (and frequently have been and are). However, it does imply that these objective categories are in principle knowable through observation; where we get it wrong is when our observations are inadequate or misinterpreted. That we are continually improving and refining our taxology suggests that we are trying to approach some standard; and this standard must be the true objective categorisation. □

Objection:
Universals are against our direct experience. We only experience individuals, and each of these are distinct beings. The idea that there is a universal, such as redness, which applies to different objects is thus counter-intuitive. •

Response:
It is important to distinguish substances from properties. Many philosophers see properties as fundamental in determining what something is (Williams, 1953; Maclean, 2014); so something is what it is because it has various properties. However, I disagree with this, believing that the form of an object (which I shall define later) is fundamental, and properties can be derived from that form. Two properties may be shared between different objects, even different types of object, because the same property can be derived from the form; for example the colour of an object can be deduced from knowledge of the band structure of the energy levels of the electrons; the band structure in turn deduced from knowledge of the individual particles and their arrangements, and this in turn deduced from the form. Much philosophical discussion of colour and other properties is, from the perspective of a physicist, very naive and ill-informed. Thus I am more concerned with substance realism rather than property realism.

With regards to the objection as it relates to substance realism, the answer is that substance is the union of two aspects: matter, which individuates the substances, and form, which is an abstract aspect and which describes what they have in common. Our experience

[8] I use the words 'being of the same *kind* or *type*' in the sense that they are individual instances of the same form. The distinction is that the word *type* refers to those species which are classed together due to identity in their form, such as ethanol molecules, while *kind* refers to those species which are classed together due to similarity, such as tigers.

is with individuals, namely the union of matter and form; however we can, and need to deduce both aspects to properly understand nature. Those who find universals unnatural are trying to apply intuition derived from observation of individual substances to the concept of form; but since individual substances are not the same as form this analogy fails. It is the intuition of these philosophers which is at fault, not the concept of universals. ○

Objection:
Suppose that there is a red apple, a red pillar box. The redness of each of these is described because they both participate in the same universal of red. However, this universal must also be red. However, where does it get its redness from? It must be by participating in something else. This leads to an infinite regression; which suggests that the concepts of universals is flawed, and thus realism is false. •

Response:
This objection rests on a particular form of realism, property realism, which I don't use. It also assumes Plato's concept of universals rather than Aristotle's. It is therefore not relevant for my own thesis. The form of an electron is not the same thing as an electron. The electron does not become an electron by participating in the form of *electronness*, but by the union of form and matter. The form of an electron does not need to be a union of two other concepts; the analogy in the objection fails because it treats form and substance as though they were the same thing, and they are not. ○

Objection:
Everything can be accounted for without universals, and physics does not depend on them. Therefore the concept of universals is superfluous, and realism should be rejected. •

Response:
The claim that *Everything can be accounted for without universals* is rather bold, and I believe disproved in the case of fundamental physics, as already discussed. Thus this objection is based on a false premise, and therefore can be disregarded. ○

Objection:
Universals are not directly observed. However, the only things we can be sure of are those which are directly observed. Therefore we cannot be sure of the existence of universals; nor is it easy to precisely say what form is. •

Response:
The issue I have with this argument is in the second premise. The first problem is that the phrase 'we are not sure of' covers a large range of positions, from as near as possible to certainty without being certain, to as near as possible to being certainly false without being certainly false. It is true that we are not certain even about many sensible objects, because of imperfect experimental precision, and also not certain about more abstract ideas. We can, however, be certain that some ideas are false. For example, we may not be certain about the precise finer details of some particular form, leaving some degree of ambiguity, but we can be certain about the broad scope (a man is not the same type of thing as an ethanol molecule). Thus the objection uses the range of the phrase 'we are not sure of'

to make it seem like because we do not have perfect surety about every tiny detail of the framework that we are in fact completely ignorant. This is a fallacy.

It is also false that the only things that we can be sure of are those which are directly observed. For example, we can be sure of the truths of mathematics (given the various definitions and axioms), even though these are not directly observed. Equally, we can be sure about the laws of physics to a certain extent, even though these are not directly observed but only deduced. It is true that we do not yet have the most fundamental theory, nor all the consequences of the theories we do have figured out, but nonetheless we do know a great deal, including ideas about how our current theories may be derived from a fundamental theory in certain limits, and (perhaps) a rough idea of what a fundamental theory might look like. We also know for sure, for example, that Newtonian physics is false (albeit a very good approximation to the true theory in certain circumstances). Of course, we know it is false because it disagrees with observation; but this is not due to direct observation, but only indirectly through deduction and then comparison. Another example is natural selection; we are sure that it occurs in nature (in part because it explains various observations), but it itself is not something that can be seen directly in nature because it is just a concept. The analogy between physics and forms is reasonable, because most of our knowledge about the details of particular forms should be derived from fundamental physics; although the general concept of form does not require modern physics to be understood. In any case, modern physics naturally leads to a concept analogous to the Aristotelian form: if there are general laws acting on various physical beings, which apply to one being as well as another, then those beings must have something in common to explain why the same law applies to each of them. This 'something in common' is a universal. For example, classical electrodynamics says that all objects of a particular mass and electric charge will act in the same way in a constant electric field. However, this requires that there are universals from which the concepts of 'mass', 'electric charge' and 'electric field' can be derived, which are instituted in all the beings in a common way. So the electric charge of one particle is the same sort of thing as the electric charge of a different particle; if this were not a universal, these would just accidentally resemble each other and so there is no reason why they would interact in the same way in an electric field. Thus the success of science seems to presuppose realism. I also do not think that modern physical theories can make sense without this concept, thus excluding all possible alternatives. So although we cannot directly observe electric charge – we only observe objects and their motions in various circumstances - we can nonetheless deduce its existence as the only possible way to explain various repeatable phenomena. ○

DEFINITION 3.2.11. **Time:**
Time is the property of material substances that describes their propensity to change; it is how one counts how many changes have occurred[9]. △

DEFINITION 3.2.12. **Space:**
Space is the property of material substances that corresponds to their extension over a region[10]. △

[9] I give here the definition as usually stated within classical philosophy, as part of my build up to a discussion of classical metaphysics and physics in the next chapters. I will revise this definition later.

[10] I give here the definition as usually stated within classical philosophy. I will revise this definition later.

DEFINITION 3.2.13. **Cosmos:**
I define the cosmos as the entirety of everything in existence, both material and immaterial. Δ

DEFINITION 3.2.14. **Universe:**
I define the universe as the entirety of everything material that exists. Δ

DEFINITION 3.2.15. **Physical law:**
I define physical law as a mathematical expression that describes the interactions between material objects. For example, if the particles in the universe may be described by the set

$$\{|e_1^a\rangle, |e_2^a\rangle, \ldots, |q_1^a\rangle, |q_2^a\rangle, \ldots, |h_1^a\rangle, |h_2^a\rangle, |\gamma_1\rangle, |\gamma_2\rangle, \ldots, |W_1^a\rangle, |W_2^a\rangle, \ldots, |g_1^a\rangle, |g_2^a\rangle, \ldots\},$$

where $|e_i^a\rangle$ refers to the six leptons, $|q_i^a\rangle$ to the six quarks, $|h_i\rangle$ to the Higgs Boson(s), $|\gamma\rangle$ to photons, $|W^a\rangle$ to the three W and Z Bosons, and $|g^a\rangle$ to the eight gluons, where each of the members of this set represents a function of position and momentum (and any other accidental properties of the particle), then the laws of physics describe the allowed ways in which these functions vary across space time as a function of the particles' position, momentum, and other properties.

In terms of individual particles, the laws of physics provide the answer to the question, 'If at time t_0 there is an electron $|e_1(p_1, E_1, x_1, s_1, \ldots)\rangle$ and an electron $|e_2(p_2, E_2, x_2, s_2, \ldots)\rangle$, where p refers to momentum, E to energy, x to location[11], and s to spin, then what is the probability that these particles will be in certain defined states at time t_1?' This I see as the primary purpose of physics. A secondary purpose is to be able to understand the fundamental building blocks from which matter is made. A tertiary purpose considers the fate of ensembles of a large number of particles in particular circumstances. Additionally, physics allows us determine the properties of a physical substance (such as its colour, hardness, brittleness, electrical conductivity and so on) from knowledge of the underlying material from which it is constructed and how that material is arranged. Finally, we can ask what are the practical benefits to mankind of these findings (for example, the invention of new technology, or better calculation of risk). Almost all questions that physicists acting as physicists are interested in answering (I would not be so bold that I would claim that all questions; but I struggle to think of any counter examples) can ultimately be expressed in terms of one or more of these five categories. The aim is to obtain a complete theoretical understanding; of course, well performed experiment is an essential part of the process of gaining such an understanding, both to check our progress and to inspire us to new ideas.

It is important to distinguish between *the* physical law, which is embedded within nature, and our knowledge of it, which is still incomplete. It is hoped that the discoveries of physicists are closely converging towards the true law. It may be that one day our knowledge converges so it is identical with *the* law, or it may be that we never find and understand the true physical law, and it does not matter for this discussion. Our knowledge of the law is a human construction which closely but not exactly approximates the true law, but it is judged against experiment which reveals the true law which is embedded in nature and which is not our own construction but something objective and external to us.

[11]Note that my notation here assumes a pseudo-Newtonian physics where one can state both the location and momentum of a particle; in quantum physics this notation would have to be modified.

I should at this point mention that there are actually two different ways of thinking about physics, and physical law. The study of physics is at least two thousand years older than Newton and Descartes, and we may ask if the concept of the laws of nature originated with their generation (or a little earlier), how people viewed physics before then. Asking questions such as this is, I think, always important, because even if we reject the answer, it does at least make us question some fundamental assumptions about how we view the world, and it is always good to identify and question our most fundamental assumptions, even if at the end of the day we still prefer to keep them. Instead of speaking of the law of nature, the scholastics tended to discuss the law of natures. Rather than one overarching equation describing everything (as Newton's second law was intended to be), every type of being was treated individually. Physics, in other words, was simply the sum of the efficient, final and formal causes of every type of being. The task of the physicist was to identify these causes. In other words, the ways in which a being can interact with other beings, and move in time, and change, is not proscribed by something external to that being, but by something internal to it. An electron behaves in a certain way because it is the nature of the electron to do so. Similarly, a lump of iron behaves in a certain way because it has certain powers which are determined by its nature. This, of course, is one reason why for Aristotle and his early medieval disciples, the studies of physics and geometry were seen as entirely separate. If physics was first and foremost about causes, then it is not obvious that these ought to be or even can be treated mathematically. On the other hand, if physics is about bodies in locomotion, and locomotion movement in space, which can be treated as a Euclidean manifold, then the link to geometry is more obvious. Now, of course, the dependency of physics on mathematics is taken for granted. But I think it is still worth asking whether we must discus a law of nature, or whether, if we can find a mathematical way of representing laws of natures, that may give an alternative and either equally or more useful perspective on physics.

It is sometimes said that the Aristotelian way of thinking about physics killed the development of physics; the adoption of the first, mechanistic, view point liberated it. Certainly there is a good historical argument for this: the point when mechanism was adopted by the likes of Galileo, Kepler and Newton was precisely the moment when there was a sudden jump forward in the understanding of physics. After all, to say that water has the power to dissolve salt is a conversation killer; to say that there is a natural law which enables water to dissolve salt leads us to try to discover the natural law and make progress.

I am not fully convinced that we can dismiss the Aristotelian position so easily. It does not necessarily kill further inquiry. When we note that water has the power to dissolve salt and (say) aluminium doesn't, we are then led to ask why water has that power when aluminium lacks it. In what way do the different forms of water and aluminium (together with the structure of salt) lead to the differences in powers? Thoughts such as this would then ask us to look at the in depth structure of water and salt, and how they would interact with each other, and from that we can proceed to develop the full theory. If done right, this might well end up in the same place as the mechanistic world view, albeit getting there in a different way. To say that water has the power to dissolve salt need not be a conversation killer; it does lead to further questions, the answer of which will lead us to more questions, and perhaps towards the most fundamental theory. The idea of having external laws of physics is easiest to start with, especially for Newtonian physics and its action at a distance for gravity which is not so easily explained in terms of powers. However, we should not so quickly dismiss the Aristotelian picture, but leave it in the back of our minds until we have written down the fundamental theory. It may be that the easiest way of describing that theory would be in terms of individual powers rather than general laws. Or it may not. We will find out when we get there. Δ

Discussion:
I am assuming that electrons, and so on, are real objects which may be described by mathematical functions. In principle, there may be other non-mathematical formalisms to describe the same objects, but the mathematical formalism is clear, simple, and well established and thus most convenient. Each particle is described by a state which I denote as $|X(x_1, x_2, \ldots)\rangle$ where X denotes the essential type of particle, and x_i the accidental properties of the particle.

It is evident from this definition (or description) that the laws of physics do not and cannot answer the principle question addressed in this work: the ultimate explanation of why things are the way they are. They provide explanations, but not the ultimate explanation. If the laws of physics are some platonic construction to which the material world happens to be bound (i.e. some form of dualism where the laws are in one realm and material reality in another), then we need to explain the bond between the material world and those laws, and why those laws and not some others. If they are a description of regularities of the universe, then they do not provide an answer: they are merely an accurate formulation of the question. If they are an expression of God's movements of otherwise inert matter, then obviously God provides a more ultimate explanation. If they are a description of the natures of beings (their actualities, potentialities and so on) then those material beings are prior to the laws. It is important then to realise that the laws of physics alone, and indeed science in general, cannot satisfactorily explain the universe, but only describe certain aspects of it. The physicist can and must contribute to framing the most accurate and precise expression of the question; he is incapable of answering it through physics alone. It may indeed be that the question cannot be satisfactorily answered by man; but, unless such an impossibility can be rigorously proved from indisputable premises, the difficulty of finding a solution should not prevent us from trying.

Note that the laws of physics are generally assumed to be outside of space and time (or transcendent), unchanging, omnipresent, eternal, infinitely precise, and unaffected by any physical objects. This assumption is required for any study or application of physics. Now it is obvious that all these properties are among the classical attributes of God. This is not a coincidence. The original founders of the concept of the 'laws of physics', the likes of Descartes, Galileo, Newton and those who inspired them, started from the premise that there was a divine lawgiver, and from that concluded that there must also be divinely given laws. Those laws must then take on some of the properties of the lawgiver. The modern atheist interpretation of physics has retained the concept of physical law, while denying the existence of God, the lawgiver, and therefore the original justification for the concept. So an important (and to my knowledge not adequately answered) question for the atheist is why do the laws of physics have those properties? Why are there laws at all? This is not a question that can be answered through physics or any of the sciences. The task of the modern physicist is to determine what the laws of physics are, and to find various practical applications of them. But the question of why cannot be answered by only addressing the question of what. We need physics to help and to guide us, but physics alone is insufficient. Some may say that there is no explanation, or the universe just happens to work that way, or it is not something we can ever know; to my mind this is unsatisfactory.

There are, I think, only six possible ways in which we can view the regularities of the universe. Firstly, that they are intrinsic to the matter in the universe (perhaps expressed in terms of powers of that matter), and fundamental in the sense that they are not derived from any higher principle beyond those powers themselves; secondly that they are intrinsic to matter but not fundamental, so that there is some immaterial principle which explains why the matter has those powers; thirdly that they are extrinsic to the matter and fundamental (i.e. an external law which imposes itself on matter and which cannot be explained in terms of anything else); fourthly that they are extrinsic to the matter and not fundamental (i.e. an external law which arises from some other non-material principle); fifthly that

the regularities are not an objective feature of reality but only arise as a convenience of our classification; sixthly that there are in fact no regularities in nature at all. The last possibility includes anything which is not contained within the first five; and obviously within each of these general classes there might be many different variations[12].

The first of these possibilities is a pure Aristotelian metaphysics; the second the Thomist variant of Aristotelian philosophy which is the basis of one form of classical theism; the third materialistic atheism; the fourth deism; the fifth I will label Humean, since one major variant along this theme was proposed by David Hume; and the sixth I shall keep nameless since the success of modern science in discovering and explaining the behaviour of matter in terms of regularities has rendered it untenable. The Humean position I also take to be ruled out by science: one has to explain why our description is so successful, and the obvious explanation is that there is objective regularity in the universe to which our description at least approximately conforms; the theory can successfully predict features of nature before they are observed, so it goes beyond a mere classification of regularities; and the theory is only consistent if the universe proceeds in the same way even in our absence (this is a requirement of the symmetry principles that form a fundamental and necessary assumption behind modern physical theory), which implies that the regularities cannot merely be a matter of our own conception and thus dependent on us. The first and second positions dominated classical and medieval thought; the third and fourth, though with their origins in the medieval period, form the foundation behind all modern philosophy except that which I have classified as Humean. The mechanical world view takes either the third or fourth position to be correct, and adds to it various further assumptions, such as that the universe is determinate (given full knowledge of the state of the universe at one point in time, one can in principle accurately compute its state at every other time), can be reduced to the sum of its parts, and that everything can be explained in terms of matter and locomotion. These assumptions are strongly challenged by quantum physics, as I shall show. It can be questioned whether materialistic atheism is coherent without these assumptions; if the source of the regularities in the universe is external to matter, is a guiding principle of matter, and indeterminate, then it can be represented as the fruits of some immaterial free will (definition 4.4.3), which is precisely what the atheist denies can exist. Aquinas also showed, and I do not believe that this has been seriously challenged, that pure Aristotelian metaphysics is untenable and collapses into Thomism. The modern world was formed from the assumption that modern science has rejected the notion of intrinsic powers; much of enlightenment philosophy mocks the Aristotelian world-view. My purpose in this work is to show that the Thomist position is consistent with modern physics, and to suggest that it provides a better match with physics than materialistic atheism or deism.

See section 6.2 for a further discussion of this point, this time in a more historical context. □

DEFINITION 3.2.16. **Mechanism:**
Mechanism is the belief that the universe is arranged in independent parts, which have no purpose in themselves, and act individually and not as part of a larger structure, and interact with each other according to fixed laws of motion which are both objective and external to them, similar to a clockwork machine. In its original form, mechanism is fully

[12]I should also mention the possibility that physics could be intrinsic to matter in some respects and extrinsic in others. For example, one might suppose that the electromagnetic and nuclear forces are intrinsic and gravity extrinsic to matter. This possibility complicates the analysis, but not by enough to disrupt the conclusions of the Thomistic world-view. I do not think that physics can be fundamental in some parts and non-fundamental in others; for any aspect of it to be non-fundamental means that it cannot be fundamental as a whole.

deterministic: if you know everything about the universe at one moment in time, then you can in principle predict what it would be like at every prior and subsequent moment in time.

> But, in order to show how I there handled this matter, I mean here to give the explication of the motion of the heart and arteries, which, as the first and most general motion observed in animals, will afford the means of readily determining what should be thought of all the rest. ... But lest those who are ignorant of the force of mathematical demonstrations and who are not accustomed to distinguish true reasons from mere verisimilitudes, should venture, without examination, to deny what has been said, I wish it to be considered that the motion which I have now explained follows as necessarily from the very arrangement of the parts, which may be observed in the heart by the eye alone, and from the heat which may be felt with the fingers, and from the nature of the blood as learned from experience, as does the motion of a clock from the power, the situation, and shape of its counterweights and wheels. (Descartes, 1637)

△

Discussion:
That some mechanistic view of the universe (not necessarily the naive mechanism adopted by the founders of classical science[13]) is true is a proposition at the heart of one of the major streams of atheism (other streams include philosophical relativism and subjectivism, which to my mind undermine any scientific view of the universe), and was probably dominant from the start of the enlightenment until the advent of quantum mechanics. Most of the major modern philosophers Bacon, Locke, Boyle, Hobbes, and so on - have been some form of mechanist; Descartes, being a dualist, cannot strictly be thought of as a mechanist, but many of his views foreshadowed this philosophy and he was close enough that we can neglect the difference. It is, in many respects, a natural accompaniment to classical mechanics, and is seriously undermined by quantum physics (though not eliminated - the many worlds interpretation of quantum mechanics tries to maintain mechanism, and the de-Broglie-Bohm hidden variables interpretation is not too far distant) and sits uncomfortably with the most natural interpretations of it. □

DEFINITION 3.2.17. **Deism:**
Deism is the belief that the universe is mechanistic, but its laws were established and put in place by God. The universe thus is capable of operating independently of God, only requiring Him to build the machine and start it off. I distinguish two forms of deism:

- **Strong Deism:** The belief that God created the universe, established its laws, but otherwise does not intervene. Knowledge of God is therefore only accessible to us through reason.
- **Weak Deism:** The belief that God created the universe, established its laws, but from time to time may break those laws to directly intervene. God is thus knowable through both reason, and observation of the miraculous.

△

Discussion:
With the dominance of mechanism as the underlying metaphysics of nature, deism became the default position of the religious, the seemingly only way to keep God in the picture given

[13]My terminology might be confusing: classical philosophy refers to the philosophy of the Greeks; classical physics (and science) refers to the physics of Galileo, Newton etc. This inconsistency in terminology is, unfortunately, standard usage, and so I am stuck with it.

Newton. Most atheistic criticisms of religion that I have seen attack some form of deism (for example, Hume's arguments largely are in response to weak deism, and are invalid against non-mechanistic views such as theism), and many modern defenders of religion (perhaps most notably Paley) have adopted some form of deism. It is not, however, a position that I intend to support in this work. □

DEFINITION 3.2.18. **Classical theism**:

I define theism as the belief that God not only created the universe, but continually upholds, maintains and sustains it. The universe cannot operate independently of God: if God ceased acting to maintain it, the universe would cease to exist. The laws of physics describe, in the general case, how God sustains the universe (although He is free to act in other ways in specific situations). Given that God is free to act as He chooses, theism, at least in its most naive form, states that nature is not deterministic. It is incompatible with and a rival to mechanism.

Classical theism adds to theism certain beliefs about God; the most important of which is divine simplicity: God is of one substance, and His substance is not composed of different parts (I shall expand on this and define the terms properly later). This should be contrasted with more modern forms of theism which dispute divine simplicity. I will define what I mean by 'God' in a later chapter. △

Discussion:

Classical Theism was the established Christian position before the enlightenment; it was (in various forms) held by the leading Christian theologians such as Aquinas and Augustine, as well as the best Muslim and Jewish philosophers. This is the position that I will attempt to defend in this work.

Note that while it is possible in deism to pit science against God as alternative explanations of an event, in theism this distinction makes no sense. In theism, science is a description of God's action; a scientific explanation is also a theological explanation. Theism does not need to postulate additional spiritual interactions in addition to the known physical ones, because in theism the physical interactions themselves describe the way God works in the world. Electromagnetism is both a physical interaction and a 'spiritual interaction' (if one absolutely must express things in such crude terms). This leaves us with the question of whether mechanism, where physical law exists independently of God, or theism, where physical law is a description of God's work, offer the best explanation of why the universe is governed as it is. This is not a question that can be answered directly from a description of that law, but we need to take a more indirect approach. □

CHAPTER 4

Aristotelian Metaphysics

I have called this chapter 'Aristotelian Metaphysics'. This title is perhaps a little misleading, and requires some further clarification. The problem is that there are almost as many Aristotles as there are commentators on him. So we have the pagan Aristotle, the Jewish Aristotle, The Muslim Aristotle, the Zoroastrian Aristotle, the Christian Aristotle and many others. Within each of these major groupings, there are, of course, many variations; within Christianity we have the Aristotle of Scotus, of Aquinas, of Anselm, and so on. All of these, of course, are inspired by the single source, and will share many definitions in common, but each school of thought developed in different ways; and in each case there were some aspects of Aristotle's writings which were rejected, and others which were adapted. Aristotle provided a framework; a way of thinking about causality, change, ethics, substance and being (among others), but each of his commentators adapted that framework so it was consistent with the particular revelations they accepted. It is wrong to overstate the differences between these approaches, but it is equally wrong to say that they are all identical. Here I will follow the interpretation of Aristotle's thought that I consider most reasonable and most useful.

In the following, my intention is to broadly follow the approach of Thomas Aquinas, since, in my view, he represented the pinnacle of Christian Aristotelian theology. This in itself means that the Aristotle I summarise is not quite the same one that the early modern philosophers attacked (and part of my intent in presenting this is so that we can re-examine the criticisms of Aristotle that began modern philosophy in the light of today's rather than sixteenth to eighteenth century physics), since fifteenth century scholasticism was not the same as thirteenth century scholasticism. However, we should see how well the criticisms match up against the best that the Aristotelians had to offer. However, I have deviated from Aquinas' metaphysics (and the metaphysics of Aquinas' leading commentators, both historically and today) in a number of ways, partly intentionally for my own later purposes, and doubtless also unintentionally in a few places because there is not complete agreement even among experts about some aspects of Thomism (and while I understand much of Aquinas' thought, I am not a professional philosopher, so naturally my knowledge is not as strong as the specialists)[1]. If you want to learn about Aquinas, then your best bet is to read both him (Aquinas, 1256, 1264, 1274; Aquinas and McDermott, 2008), Aristotle (Aristotle and Barnes, 2014), and the Jewish and Muslim thinkers he drew on (Al-Farbi, Averroes, Avicenna and Maimonides). If you want to learn about how he is usually interpreted today, then the twentieth century manuals are one place to start (Sullivan, 1957; Mercier, 1900; Baschab, 1937; Garrigou-Lagrange, 1950; Maritain, 1962; Chesterton, 1933); more modern

[1]For example, there is a distinction in classical Aristotelian thought between natural substances and artefacts which I will not dwell on, instead glibly discussing forms of tables and the like as though the same word 'form' applied in the same way to both tables and trees; something which purer Aristotelian philosophers than me would consider a dangerous oversimplification. Also form in Aristotelian metaphysics generally refers to species rather than individuals, while I will define form so that individuals can also have their own forms; so, for example, I will say that there is a particular form of newts in general, and that each individual species of Newt has their own form which inherits the form of the general newt plus some additional restrictions; and each individual newt has a particular form which inherits the form of the species plus adding some further restrictions – again, this is a deviation from mainstream thought as I understand it.

writers include Oderberg, Gilson, Feser and plenty of others who can well describe the conventional approach (Feser, 2014, 2009; Oderberg, 2008; Gilson, 1994, 2003; McInerny, 1982; Stump, 2005; Kerr, 2009; Geisler, 2003) – although you should be warned that many of the older works assume a familiarity that comes from several years of study of Aristotle.

My intention, however, is to find a philosophy that fits modern physics. That means that I will take from Aristotle and Aquinas what I need, and discard what doesn't work, while also drawing on ideas from other sources and my own thoughts as they are helpful. My goal is the truth, not to be faithful to dead men, so I will only accept Aristotle's writings to the extent that they are useful in providing a true description of reality, and I will reject whatever is in them that is false. Of course, how much is useful and how much is false is something we still have to determine; most modern philosophers will say that there is far more that is useless in Aristotle than there is to help us. But let us not prejudge the man before we hear what he has to say. My intention is not to provide historical commentary, but to find a philosophy that works today; and while older philosophies can provide an underlying foundation and framework which might be a useful inspiration, obviously they will need modification and perhaps substantial alteration before being fit for the contemporary view of physics.

4.1. Introduction

Aristotle's metaphysics was (as I read it) first and foremost an attempt to solve the problem of change and constancy. It seems from everyday experience that certain objects remain the same over long periods of time. It also seems that certain objects change. The difficulty is that we also have objects which both change and yet remain the same type of object. For example, an infant grows into an adult; there is a change, but we usually ascribe them both to be the same individual. Or we can consider a man who lies down one moment and then sits. There is a change in the body posture, but one and the same man. The question is how best to express the idea that things can both remain constant in their type of being, and yet exhibit change in certain aspects of their being.

Before Aristotle, two philosophers had offered opinions on this question[2]. Parmenides responded to the problem of change by supposing that change was impossible. To him, for a being to change involved it going out of existence and something else coming into existence. But being itself could not cause the change, because during the process of change it ceased to exist. Now the opposite of being is non-being; so if being did not cause the change, then the cause of the change must be non-being. However, non-being cannot do anything. Thus neither being nor non-being can cause change, argued this over-simplified caricature of Parmenides, and the only reasonable conclusion is that change is impossible. Thus nothing can change, and anything which seems to us to change must be an illusion.

Not so, responded the traditional caricature of Heraclitus. There is another option, which is there is no such thing as being. Everything is in constant flux; you cannot step into the same river twice, because the individual flow of water has moved on. Each man is eating, absorbing nutrients, breathing, and emitting waste. He absorbs matter and expels matter. No man is exactly the same in any two moments. Therefore there is no such thing as being, but only change and coming to be. Instead of saying there is no change in the universe, instead we can say that there is no constancy.

These positions might seem to us to be a little far fetched; in part that is because of my rather poor presentation of them. However, they have surfaced in different forms down the centuries. I can, for example, see a little of Parmenides in a mechanistic position inspired by Descartes. In this position, matter just is; it cannot be created or destroyed, so that is one form of change denied immediately. Instead, matter continues as it is unless acted on by an external force. The inertial state does not change. The question is, of course,

[2]Please note that I am giving the simple and slightly exaggerated version that appears in introductions to Aristotle; it is useful to show the two extremes before considering Aristotle's position. In reality, both Parmenides and Heraclitus were more nuanced than this (Osborne, 2004; Waterfield, 2000).

how do we understand change. If we mean change to imply change in momentum, then the principle of inertia suggests that there is no change under normal circumstances. This is not, of course, Parmenides' position, because there are still external forces. According to Descartes, matter is just extension. The only change is change in location and velocity: everything else in Descartes' thought remains constant. Particles can move around, bump into each other, arrange themselves in various ways and come apart, but the type of matter and type of being does not exhibit change. Everything apart from this has been moved from the world of the physical to the world of the mental, and in Cartesian dualism there is an impenetrable barrier between the two. Descartes' position is thus part of the way towards Parmenides. However, in terms of location, everything is in motion. Things are always moving from one place to another; there is no such thing as being at rest, because what is at rest to one observer appears to be moving to another (the second observer is travelling at constant velocity). So for locomotion, Descartes sees no constancy: he adopts Heraclitus' position. We can, of course, go another way by combining Cartesian principles with Einstein's physics. In general relativity, the force of gravity is in some sense absorbed into the definition of what it means to stay constant. We can think of the principle of inertia as saying that things continue in straight lines unless acted on by an external force. The question is, what do we mean by a straight line? In geometries other than Euclidean, the answer is not as simple as we might be used to. What makes a straight line depends on the curvature of space; the equivalent to a straight line in a curved space is known as a geodesic. The effect of gravity is to curve space time. In other words, in general relativity, it is not the planet in an elliptical or circular orbit whose path is curved, but the path of the planet is straight but space itself is curved. For example, one can consider moving in a straight line on the surface of the sphere (in this case 'straight' is defined as always being tangential to the surface of the sphere at that local point): one will eventually get back to where one starts. If we define change as movement from one geodesic to another, then one could say in a system with gravity and nothing else there is no change because things naturally remain on geodesics. In other words, in some sense, we have reached Parmenides' changeless universe (if one accepts this definition of change; I recognise that most people won't).

Hume, on the other hand, I think if his principles are taken to their conclusion, tends towards the direction of Heraclitus. For Hume is an empiricist, and for him only what we observe or can identify from mathematical axioms is worthy of consideration. Now we recognise things changing: after all, every act of observation involves change: light bouncing off a surface, interacting with our eyes and so on. However, being is not something that we can observe; it is something that we infer. I do not observe that my wife remains the same woman that I observed yesterday, I just infer it based on the appearance of continuation. But Hume rejects that we can know anything for certain from inductive arguments such as this. What if my wife's body eased out of existence during the night, only to be replaced by something apparently identical? Would she be the same woman? Hume believed that we cannot have any metaphysical knowledge, and being is a metaphysical category. We observe change, but do not observe being. Hume could, I think, slip easily into the position that there is no being, but only change.

Kant, however, takes a compromise position. He divides the world into the world of ideas and the world of reality; and contends that the two cannot match with each other. In the world of ideas, he has constancy, since our image of a particular cat remains the same; in the world of reality, however, there is anarchy and confusion: it is basically Hume's world, where anything can in principle happen, and there are no metaphysical concepts such as being.

So while no modern philosopher would adopt exactly Parmenides' or Heraclitus' position, the two of them are sitting around like ghosts in the background. The one thing all these philosophers have in common is their rejection of, or perhaps failure to consider,

Aristotle's solution. Philosophers tend to echo Parmenides in certain respects, and Heraclitus in others, but with each kept within their well defined and separate boxes. And this is, I think, not surprising. For the only reasonable positions are some variation of either Parmenides, Heraclitus or Aristotle's thought: we can combine them to say that in certain areas of physics or certain times Parmenides got it right and in other areas or at other times Heraclitus got it right, but it is always a combination of the two, unless we go for Aristotle's solution. And the modern world has rejected Aristotle.

The Aristotle I'm presenting here (who isn't quite the same as the historical Aristotle - I am rephrasing and simplifying - but he is close enough) believed that both Parmenides and Heraclitus had made the same mistake. They assumed that each being only had one possible mode of being; the choice is between being and non-being. Aristotle added to this another mode of existence: potential being. Change was not something ceasing to exist and something else coming into existence, but actual being and potential being changing their modes of existence.[3]

So it was said that a man had to have his matter arranged in precisely that way, or he was not the same man. In reality, however, there are numerous different ways we can be the same man. So we can define any individual being, say X, as something that exists in one of a larger number of possible states, say X_1, X_2, X_3, ... So to say that something is X means that it is either X_1 or X_2 or X_3 or .... Both Parmenides and Heraclitus (and a number of more modern philosophers[4]) implicitly said that there was only X_1; it could only be X if it was in that one state. But this is false. So, for example X_1 could be the reclining man, X_2 the sitting man, and so on. At any given moment the being, if it exists, must be in one, and only one[5] of these states. That state is known as the actual state. What of the other states, X_2, X_3 and so on? These can still be said to exist, since they are present within the concept, or form, of X. However, they do not exist actually; instead we say that they exist potentially. So Parmenides made the mistake of saying that there was only being and non-being; in practice, if Aristotle is right, there are three possibilities: actual being, potential being, and non-being. Change is not being turning into non being and non-being turning into being, but actual being becoming potential being and potential being becoming actual; this process is known as actualisation.

The potentia lead to several additional concepts. Firstly, they tell us what change is possible and what change is impossible. They tell us that something which is actually X_1 can (in the right conditions) turn into X_2, but not Y_2. In other words, they not only enable change, but can illustrate why some types of change occur and others don't. The also give us an idea of what is meant by continuation; or what exactly it is that continues when change occurs in a being: it is whatever each of the states have in common, for there will be some things they each have in common, and some things they don't. Each of the X_i are abstract ideas; they might indicate a certain arrangement of matter, or perhaps something more sophisticated than that, but they each describe one particular state of being. That

---

[3]I should note that Aristotelian philosophers have distinguished between active and passive potentials. An active potential refers to the capacity to induce change in other beings. A passive potential refers to the capacity to be changed by something else. In this work, when I refer to a potential, I will mean passive potency unless it is otherwise clear from the context.

[4]Kant, for example, if I understood him right made the same error, saying that our idea of something existed only in one state X_1; but in practice that could never match up precisely with reality, so our idea is distinct from reality. However, in practice, our conception of things need not be so simple. Our conception of X, if we are smarter than Kant was, includes not just X_1 but the entire set of possible states of X. Although I have treated the different states as discrete entities for ease of explanation, in practice they form a continuum. So the actual being is one point in that continuum. That means that the same being exists in reality as in our concept; the two do coincide. Kant regarded our conceptions as precise but inaccurate; I would rather say that they ought to be regarded as accurate but imprecise.

[5]That it can be in only one state seems obvious, but this is a point which needs further discussion; I will return to it because it is not as simple as it seems.

means that the set of all the X_i is also an abstract concept, but one that describes almost everything that it means to be X. I label the set of all the X_i as the form of X.

The one missing ingredient is the matter itself. Form is just an abstract concept and is common to all beings of the same kind; matter is what makes the being tangible and individuates different beings of the same kind. Thus being, or substance, is a union of form and matter. This tells us that the arrangement of matter is just as important in understanding a being's properties as the underlying matter itself.

Objection:
In the above analysis, change within a substance X is described as the actualisation of potentia. However, this does not explain change from one substance to another, for example hydrogen burning in oxygen to become water. In this case, we once again have to see being becoming non-being and non-being becoming being, which is impossible. •

Response:
Aristotle solved this dilemma by introducing one more possible state of being: virtual existence. An isolated hydrogen atom can exist in the states H_1, H_2, and so on. An isolated oxygen atom in states O_1, O_2 and so on (in practice, of course, the isolated atoms are unstable except when as part of a molecule, but for simplicity I shall neglect that). Water can be in states W_1, W_2, W_3 and so on. Two H_1s and one O_1 do not by themselves make one W_1. That is obvious: the properties of water are very different from the properties of the two gases. They have to be absorbed into the water molecule. However, the hydrogen and oxygen do have some sort of continuation as part of the water molecule. This is described as saying that they exist virtually inside the water molecule; in the same way the water molecule exists virtually inside the isolated hydrogen and oxygen. Thus substantial change can be described as virtual existence becoming actual existence, and actual existence becoming virtual existence. This is, of course, similar to how we treated the potentia. The difference is that potentia refer to different types of being of the same substance, while virtual states relate to change of substance.

I will later distinguish between these two cases by saying that a change within a substance can be accomplished by a continuous change; a change between substances requires a discontinuous change in the arrangement of matter and the other quantities that describe a state. ○

4.2. The Premises of Aristotelian Metaphysics

As I read it, Aristotle's metaphysics is built on the premises below. I will provide definitions later in the chapter.

(1) Material objects can be placed together into various sets, each set representing objects of the same fundamental type.
(2) These types are intrinsic to the universe; they are not simply a matter of human construction or convenience.
(3) There is a unifying principle, form, which is an abstract quantity underlying the similarity between the objects.
(4) There is a distinguishing principle, matter, which is not abstract and differentiates between objects of the same form.
(5) A material substance is the union of form and matter.
(6) A material substance is not simply the sum of its parts. When two substances are compounded, the result is a third, distinct, substance with its own form which cannot be reduced to the forms of its constituents.
(7) We can grasp and understand form in the intellect. In this way, we can gain an understanding of the world around us.

(8) We can neither grasp nor understand matter in the intellect. In this way, our mental understanding is not a complete representation of the world around us.
(9) Within each form, there are numerous states, called potentia, which represent the different ways or structures through which that form can be manifested.
(10) Change (or motion) is defined as the movement from one potentia to another (often, but not always, without any difference in the underlying matter).
(11) At any one time, each particular object will exist in one particular potentia; this state is known as being actual.
(12) Another possible type of change is the creation or destruction of material substances. It is not impossible that even the most fundamental building blocks of matter can be created and destroyed.
(13) Every motion from one potentia to another, or creation or destruction of a substance, is caused by at least one substance already existing (i.e. not the destination potentia); this cause might be the same object in its initial state, or it might be that plus one or more external causes.
(14) There is a limited number of effects which a potentia can cause; these are implicit in the definition of the form. The effect is thus in some way contained within the definition of the cause.
(15) Properties can be understood from knowledge of the form.
(16) Properties can be either essential or accidental to the form. Not every property is essential to the form.

4.3. Being and Matter

DEFINITION 4.3.1. **Being:**
A being is anything which either does or possibly may enjoy a stable or semi-stable existence[6] (I exclude those things which are in a state of instantaneous decay, or in transition between two types of being). This may be a living creature, non-living object, immaterial concept, or anything else. △

DEFINITION 4.3.2. **Matter:**
Matter is that from which all sensible (i.e. what we perceive with our senses, or can in principle perceive with our senses) objects are constructed and which gives them concrete existence; it is the actualising principle for the substance. Matter is the basic building block of all material objects; we would today identify it with the physical presence of quark, lepton, Higgs and gauge Boson fields (along with any the fields for other fundamental particles not yet confirmed to exist).

Matter is also the principle that individuates objects with the same form. Each object of a particular form has its own individual matter: no electron is shared between two different tables, for example. Any object which is incorporated within another object (as the toe is part of the foot) will have its matter as part of the matter of the larger object: all of the matter within the toe is part of the foot; not all the matter in the foot is part of the toe. △

DEFINITION 4.3.3. **Form:**
In its most naive representation, form is a description of how the matter is arranged to create a particular substance. More accurately, form describes all the possible ways the

[6] See 4.5.1 for my definition of existence.

matter might be arranged so that an object is of a particular substance; and it will also specify everything apart from the arrangement of matter that contributes to the structure and properties of the object (if there is anything beyond the matter and its arrangement that does so). It is the set of all possible states (potentia) which could describe a being or type of being. The form of a cat describes every possible way that the underlying matter could be arranged that produces a cat rather than a dog or anything else that is not a cat.[7] Form by itself is an abstract concept; but it can be instituted in matter or an intellect and then it becomes joined to the physical world. △

Discussion:
When considering stable compound objects, such as water, note that the form is not simply the form of the parts of the object. The form of water is not the form of oxygen plus the form of hydrogen, but something entirely different (thus water can have very different properties to its constituent elements). But is the matter just the sum of its parts? The naive answer is yes; the matter that contributes to the hydrogen and oxygen atoms is the same matter that actualises the form of the water molecule. More precisely, however, the atoms are held together by electromagnetic forces mediated by the exchange of photons, and these virtual photons are present as part of the water but not (at least not in the same way – recall that form does not represent a single state but a set of possible states) present as part of the atoms. But this is a small quibble; to a very good approximation matter is what is preserved when elementary particles form a compound, and form is what changes when two hydrogen molecules and one oxygen molecule combine to form two water molecules. □

DEFINITION 4.3.4. **Properties:**
Properties are the observable qualities of substances: length, mass, colour, weight, hardness, electrical conductivity and so on. These can in principle be derived from knowledge of the form.

Note that this approach is different from some other approaches to metaphysics, which describe being as a bundle of properties. In the approach I describe here, being is more closely related to the form; and from knowledge of the form we can understand what properties a being can have. Of course, the properties are what we directly observe and sense. Form needs to be deduced by either induction or through a process of falsification (or perhaps some other means), and is firstly more fundamental, and secondly unifies and explains the various properties of a being. △

Discussion:
For example, diamond and graphite are constructed from the same matter, but as the carbon atoms are arranged differently, they have very different properties. These different properties ultimately arise because of the different forms of diamond and graphite, with these forms including a description of the different crystal structures of the two materials (although, of course, each would have numerous potentia since there are minute possible variations even in a diamond crystal, such as the vibrations of the atoms, or the valance electrons moving from one energy band to another); from these forms we can construct the properties of each substance: its mass, extension, brittleness, shape, colour, opacity, density and so on can all be calculated from the different structures of the material; and the structure is part of the form. Each object of a particular type shares the same form: two different (but identical in structure) tables have different matter, but the same form. Form can also be applied (in a slightly different but still related sense) to describe and distinguish abstract concepts.

[7] I will discuss the definition of form in more detail later.

As another example, a cat may be reduced to a certain number of fundamental building blocks, principally protons, neutrons, and electrons[8]; however these particles by themselves do not make a cat. *Form* is that which when combined with the underlying matter creates a material object of a particular kind. Note that it is form, not matter, that determines what something is. One could, in principle (though not in practice), take a dog or a horse (or anything else), extract all the individual protons, neutrons and electrons, put in other protons, neutrons and electrons in their place, and the creature would be of the same form, and therefore the same dog or horse[9].

Form is an abstract concept which by itself does not make up the substance but, when united with matter, combines to produce the being. In nature, it is only manifest as it appears in various physical objects, but it can also be grasped by an intellect. The external attributes of an object: colour, size, shape and so on, should not be identified with the form, but follow as a consequence of the form. Not every conceivable being can exist in nature, only those with well defined form; though not every possible form is actualised in reality. The form describes not only what a substance is, but what it can be while remaining the same substance; what properties and aspects of the substance can change, and how they can change, and which properties are essential to it. The form is more than just the sum of its parts, and new surprising features can arise when we combine two simpler substances. Therefore, while there is much use in considering the various parts of the being on their own, to do so will rarely give us a complete picture of the whole of the being: we must also consider how the various parts interact with each other. The form of the object also tells us what it may decay into; and the resultant objects (and all that follow from them) reside either potentially or virtually within the form.

Because the form describes many different actualities, it can equivalently be thought of not as a positive statement – this is what a cat is – but as a series of negative statements – this is what a cat is not – which are complete enough to eliminate everything that is not a cat. In this way we can have inheritance from a higher form (the form of an animal) through to lower forms (the species of cat, then the breeds of cats, then individual cats). So the species of the cat inherits all the restrictions pertaining to animals (it is not an immaterial object, it is not a non-living object – i.e. it is not an object which does not contain some complex molecule which carries information and the ability to make copies of itself, it is not an object incapable of self-locomotion and so on), and individual cats inherit the restrictions related to the species of cat (it is not an animal based on something other than DNA, it's DNA is not that of a pig, or a giraffe, or a mosquito and so on), plus some additional restrictions differentiating one cat from another. This overuse of double negatives is clumsy grammatically, but I think the best way to think about the concept of individualisation. An equivalent approach is to say that there is a boundary in the space of all possible beings within which is enclosed a space where all beings of a particular form reside, and within that space we can draw additional boundaries each enclosing a smaller region describing a less general form, and so on until the point where we cannot draw another boundary without splitting the possible arrangements of an individual into two or more of these regions.

This definition of form has its problems; for one thing it seems to be circular, in that the definition depends on the concept of substance while the definition of substance depends on the concept of form. Though we can intuitively distinguish between different types of being – cats, dogs, electrons, kiwi fruits and so on – this definition is not entirely satisfying once we get beyond the intuitive level. How do we know what are the different potentia within a form, and what is a different type of substance altogether? This objection is not insurmountable – the definition can be improved, and I will give a more precise definition

[8]There will also be the gauge Bosons holding the particles together, and some virtual pion/anti-pion pairs – as an example – created for short periods of time by quantum processes.

[9]This point is controversial, and I expect most Thomists to disagree. I think, however, given the indistinguishably of of the different fundamental particles it is better to accept it.

of form later, that does not depend on the idea of substance, but that will have to wait until I have the underlying foundations for my definition in place. Others, of course, have more carefully defined the concept of form in other ways. However, this intuitive definition will suffice for our present purpose of giving a quick overview of Aristotelian Metaphysics. □

DEFINITION 4.3.5. **Substance**:
Substance is the union between form and matter (for material beings; for immaterial beings it just refers to the form). The form and matter are enough to describe a sensible object and distinguish it from all other sensible objects. Substance can describe any stable or meta-stable[10] being or collection of beings, whether it is indivisible or assembled from simpler parts. This definition is in a different sense as to how it is applied in modern usage. More usually today, we think of substances as indivisible or simpler beings: we would not classify a star as a substance, but a compressed hydrogen gas would be a substance; in the classical metaphysics I use here both stars and hydrogen are substances in their own right, with the star composed of hydrogen gas (among other things). △

DEFINITION 4.3.6. **Material**:
A material object is a being that contains matter (or energy). It (or objects composed of it if it is too small) is in principle perceivable by our senses. △

DEFINITION 4.3.7. **Immaterial**:
Something immaterial contains no matter, so is composed of form alone. For example, abstract concepts, mathematical or logical propositions, moral ideals such as justice, courage, and so on are immaterial. △

Discussion:
There cannot be a being which is matter alone, since in every material object the matter must be arranged in some fashion, and therefore can have form. Form, however, can exist without matter. The form of a triangle exists whether or not there are any triangles in nature. A first order differential equation has a particular form (where the word is here used in an different but analogous sense to the definition above) which distinguishes it from a second order or ordinary quadratic equation.

Form can be said to exist in three distinct modes, either by being conjoined to matter actually (for example, the form of the table exists within every table); or the states which constitute the form being present in matter potentially (such as a child is potentially an adult in the sense that it may transform into the adult; and thus in some senses the form of the adult is present within the child, in the same way that the form of a rock isn't, since a

[10] I use 'stable' in a slightly loose sense, meaning that the macroscopic properties on average are unchanging in time or only change very gradually in time according to the beings internal final causes. So neither small oscillations in the positions of atoms nor the gradual development of a living organism count against this definition of stability. Recall that form does not represent a single state, but a set of potentia, so movement from one of these potentia to another does not qualify as a change of substance or a breakdown of this stability. A continuous change such as the vibration of an atom represents a movement from one potentia to another, so does not imply a change in substance. An atom being ejected from the object would qualify as a change of the substance of the individual object, since this corresponds to a discontinuous change in the form. By stable I mean that changes such as this do not occur except through intervention of some external force; by meta-stable I mean that changes like this occur infrequently.

child cannot directly become a rock) or virtually; or within an intellect (I will later define an intellect as that which is capable of grasping abstract concepts such as forms without having the matter take on those forms). An intellect can thus understand a dog by grasping the form of dogs in general, and then the form of the particular dog in front of it. It is not thus necessary that any examples of the form exist in practice for the form to still have meaning; there are no unicorns or dragons, but we can still discuss, define, and grasp the form of an unicorn or dragon. Forms are also then (in general; there may be specific exceptions) universal, in that they are not restricted to any particular time, place, matter or intellect.

However, it should not be thought on this account that forms of actual objects are subjective. The form of an electron is determined by electrons in general; we do not define an electron as a particle with charge -1, spin $1/2$ and a rest mass of around $0.511\text{Mev}/c^2$[11] and then happen to find such a particle in our experiments; we find such a particle and then discover its properties, which in turn allow us to compare against what is expected from possible forms of electrons. More precisely, we can construct forms of every logical possible particle, and then give the name 'electron' to the particle we observe in nature whose properties are identical to the properties of that particular form. From the hypothetical forms we keep all those which are consistent with the experimental bounds on the electron's properties, stating that one of those forms we will be able to narrow it down more as the experimental precision improves – corresponds to the physical electron, and discard all the rest (except, obviously, those which could correspond to other types of physical particle). For fictional beings, we may choose the form as we like; for real beings we observe the existence of the kind and compare those observations against the form; the better our observations, the more precisely our understanding of the form will align with its true nature. The form of physical objects is not an arbitrary choice or convention that we happen to make, but it is objective and (in principle) empirically discoverable.

Forms also apply for species and individuals within that species: *dog* has a particular form which is shared by each example of a dog and which distinguishes it from the form of a *cat* (for example); in individual breed of dog has all the attributes given to dogs in general, plus others which distinguish it from other breeds; while each dog itself contains the form of its breed plus additional attributes which distinguish it from other members of that breed. However, the organs of the dog do not have their own forms, since they do not have an independent existence as they cannot be separated from the dog for long periods while remaining alive; instead we say that these are parts which together combine to make a complete dog.

Because it relies on this concept of forms, Aristotelian metaphysics is clearly realist. Many modern philosophers deny that there are these universal objective forms; in fact the denial of objective forms, or, equivalently, formal causes (or more specifically the possibility of potentiality; the idea that a being can exist in numerous different states) is (along with the denial of final causes) one of the basic foundations of enlightenment philosophy, and the philosophies of Descartes, Locke, Hume, Kant, Rousseau, Mills, Marx, and so on implicitly or explicitly rely on this denial. If they were wrong in this regard, as I believe that they were, it calls into question their entire systems; while they may have been right about many things, their arguments for those conclusions, being based on this false premise, need to be rejected and better arguments put in their place.

I do not think, however, that it is possible to conceive of physics without forms. A π^+ particle, for example, is composed of the arrangement of up quark and anti-down quark, together with the gluons binding them together, to form a spin zero particle. It has a certain

[11] A Mev/c^2 is a unit of mass frequently used for convenience by particle physicists. It is equal to roughly $0.000000000000000000000000001782662$ or 1.782662×10^{-27} grams. That number alone should be enough to explain why particle physicists prefer to use Mev/c^2 rather than grams. I should also say that the number 0.511 does depend on our convention, the choice of units, but the mass is still objective – for example the ratio of the electron to proton rest mass, roughly 0.000545, is dimensionless and the same no matter our choice of units.

mass, electric charge, parity, and every π^+ throughout the universe is the same (once we define the π^+ as a member of a set containing multiple states rather than a set containing just a single state; different π^+ particles might be in a limited number of different excited states, but these are all part of the same form). These properties are built into the laws of physics. You cannot have QCD[12] without having the possibility for pions, and the nature of these pions is fully determined. It does not matter whether there are any pions in existence at a particular time, the form of the pion is as universal as the laws which describe it. □

DEFINITION 4.3.8. **Essence:**
The essence of a being refers to a complete description of its form, matter and causes. This is enough to fully understand what a being is. In a second sense, it can refer to those features of the being which are common to all members of that kind of being (i.e. which all are members of the same set described by the form). △

DEFINITION 4.3.9. **Essential:**
Essential properties are those properties which a being, if it lacked, would cease to be that type of being. A form consists of numerous possible states X_1, X_2, X_3, ...; an essential property is one which is possessed by all of these states. △

DEFINITION 4.3.10. **Accident:**
An accidental property of a substance is one which if changed (within a certain specified range) would not make it change its type of being; for example, the colour of a table is accidental (while the colour of a red table is essential – at least up to the range of colours described by *red*), or whether is is made of wood, or plastic, or metal (but it cannot be made of jelly, as that would conflict with its essential property of being solid). In other words, if we derive the properties from the form, then different arrangements of matter that are part of that form may have different accidental properties. So, for a form with possible states X_1, X_2 and X_3 a property shared by X_1 and X_2 but not X_3 is accidental.

A second sense of *accident*, often used in classical philosophy, refers to something which subsists in something else; in other words something whose existence depends on the existence of something else. A substance exists in its own right. An accident only exists inasmuch as that substance does. Examples of such accidents include colour, taste, texture and so on. However, I will generally use the word *property* or *attribute* rather than *accident* to describe such things, to avoid confusion with the notion of accidental properties defined in the previous paragraph.

Note that this is not the common modern sense of the word where accident means (for example) something coincidental or unintentional. In this work, if I use the word 'accident' without further qualification, I will mean it in the sense of a property not shared by all the potentia within the form. △

Discussion:
Essence is clearly related to form, but is not identical to form; it is essential that a table

[12]Quantum Chromodynamics, our theory of the strong nuclear force, is the theory that describes the formation and properties of pions.

is constructed from matter while form is immaterial. Aristotle identified essence with substance. □

DEFINITION 4.3.11. **Natural**:
I will understand *nature* in two senses. In the first sense, I will denote it as the entirety of those beings whose form must be conjoined to matter for it to exist and which has a presence within space-time. When I speak of nature as whole, I will mean it in this sense.

The second sense means in accordance with a being's own inherent form and inherent tendencies; that is to say that it is a member of the set of states that make up the form of the being, and one that is consistent with its own inherent tendencies. For example, an acorn and oak tree are both natural, because acorns have the tendency to develop into oak trees, and both correspond to states which are part of the form of the species *oak*. However, a table is not natural, since oak trees do not have the tendency to become tables, and nor is a table an element of the form of the species *oak*. To construct a table, therefore, we have to work against its nature, and thus we call a table artificial, since the word *artificial* means to be contrary to the form and inherent tendencies of the being. One can take the DNA of *oak* and, combined with knowledge of physical processes, conclude the existence of oak trees. One cannot from these premises alone conclude the existence of tables. Similarly, when the various chemicals are part of a plant, they are natural, since they develop as part of the form of the plant from its inherent tendencies. But once we cut off twigs and roots of the plant, and boil them as medicine, then they cease to be natural because it is not an inherent tendency of the plant to be used as medicine, nor is the medicine part of the plant's form. Such medicine is just as much an artefact (something which is artificial) as a medicine produced through chemical manufacturing. When I discuss the nature of a particular being, I will generally use it in this sense. △

Discussion:
This distinction in the word 'nature' is part of the difference between a mechanistic and Aristotelian understanding of physics. When a mechanist discusses the 'law of nature', they mean nature in the first sense; something which guides the universe as a whole which has little care for the individual details the individual particles which are bound to them. When an Aristotelian discusses the 'laws of natures,' they mean *nature* in this second sense, i.e. the sum of the inherent tendencies of particular beings, with little reference to the wider universe, except those parts immediately interacting with them.

There is another common use of the word 'nature', which means *not domesticated or constructed by mankind*. Someone who uses this definition will call wild flowers, rabbits and polar bears natural, and buildings and philosophies artificial. So when, for example, I might say that it is unnatural for a woman to eat her man after conception, people who use this definition will complain 'It is perfectly natural: spiders do it all the time.' However, I mean that it is against the inherent tendencies for women that they should eat their sexual partners; rather the natural tendency is for them to complete the task they have started, by together working to raise the child. What spiders get up to in their own private lives is completely irrelevant to what is natural for men and women (indeed it is of limited use in determining what is natural for spiders), since spiders have their own inherent tendencies, and mankind theirs. I do not think that this definition is useful for my purposes, and so will not use it in this work. To respond to my work by saying 'Polar bears do it, so it must be natural' is to commit the fallacy of confusing two contradictory definitions of the same word. □

DEFINITION 4.3.12. **Supernatural:**
I define a supernatural being as one whose form need not be conjoined to matter. It is opposed to the first sense of the word *nature*, but not necessarily the second. For example, mathematics, physical law, logical propositions, and forms as abstract descriptions are supernatural. △

Discussion:
This definition of natural and supernatural is obviously related to material and immaterial; the difference is that material refers to individual objects which exist, while natural refers to a type of being which need not exist in practice. For example, a particular dog is a material being, while dragons (in general) are natural but immaterial, as there are no dragons in existence, yet if there were they would be part of nature.

This definition of supernatural is very different from the popular conception, of ghosts, pixies, things that go bump in the night and so on. The modern mind, of course, does not accept the existence of forms; once those have been removed, then the modern conception is all that remains. However, I believe that my definition here is closer to the original meaning of the term.

Note the relation between the supernatural and the intellect; supernatural beings are those which are accessible to the intellect but are not accessible to the senses, even though the senses provide us with enough images to give us enough hints to contemplate the supernatural. For example, the presence of triangular objects in nature inspires us to contemplate the supernatural concept of triangularity; without the natural triangles we would not be able to construct the supernatural concept of triangularity. □

Comment:
We group beings into categories or classes because of certain similarities between them. We identify similarity through comparison. We can compare objects with each other in several ways. There may, for example (I do not claim that this list is complete), be similarity in material; similarity in form (or the objects are two individual examples of a more general form); similarity in function; similarity in origin; or similarity in properties (for example, we may compare the colour of two different objects). We may also compare things analogously, by saying that one object is like another in certain respects while not sharing the same form, matter, origin or effect.

In other words, when we think about categories of beings, we need not only think about similarity in form. Every form gives us a category, but not every category is based on form.

For example, let us think about livers. With the definitions used here, there is not a single form of the liver; the liver is rather part of the form of an individual man. We cannot just ontologically remove livers from men and draw a line around them without that line cutting through the form of men. Livers can be compared in terms of their purpose rather than their form; since each liver shares a similar function within each man. In terms of arrangement of matter, the comparison is one of analogy rather than form; the arrangement of one man's liver is like that of another man's liver in that there is a great deal of resemblance between them in terms of the macroscopic structures. However, livers are not of the same form because at the molecular level they each contain different patterns of DNA, and since the liver is not the only thing that shares that DNA (a man's heart has the same DNA as his liver, excluding epigenetic distinctions and the occasional mutation) we cannot just lump all livers into one form. A form is a distinct region in the space of all possible beings, and an object cannot be the part of two different forms which are not ancestors or descendants of each other; therefore a liver cannot be part of the form of a man and part of the form of livers since not all of the man is a liver.

So we might define a liver as being part of an animal body which performs a certain set of functions (it does not matter if we do not know what all those functions are: reality is not constrained by our lack of knowledge of it); this would represent categorisation primarily in terms of the final cause (although we have additionally specified that it is part of the form of animals). Alternatively, we might define the liver in terms of an analogy; that livers are objects with certain similarities in structure. The analogous definition is, however, weaker because it does not correspond to something related to the object's essence; it is merely a linguistic convenience. The final cause of a being is, however, part of what that being is. Categorisation in terms of function, like categorisation in terms of form, material, or efficient cause, is not therefore a matter of language and only a human invention, but instead truly reflects something in nature. □

4.4. Mind and Life

DEFINITION 4.4.1. **Mind**:
Mind is the union of self-awareness, intellect, will, indeterminacy (or freedom), and perception. △

DEFINITION 4.4.2. **Intellect**:
An intellect is a substance which contains the ability to grasp abstract concepts, including logical propositions and forms, without changing its own substance. △

Discussion:
Form may manifest itself either in matter (for example, the squirrel in my garden manifests the form of a squirrel) either actually, potentially or virtually, or in a second sense in an intellect. When matter actually takes on the form of a squirrel, it becomes a squirrel; a change in substance. When it might change into a squirrel, we say that it has the form of a squirrel potentially or virtually. But these are not the only ways in which the form of a squirrel may manifest itself; I may think of a squirrel even though I am neither a squirrel nor likely to turn into one. When I think of a squirrel, what I think of is its form (I may, of course, be inaccurate or imprecise when I think of the form, but that does not affect the concept. I am still thinking of a squirrel, even if it is not precisely the squirrel in front of me). Thus there is another way in which the form may be grasped, which is neither actual, potential nor virtual. I define an intellect as anything that can take on a form in a way that is not actual, potential or virtual. Thus form must contained within something either actually, potentially, virtually or intellectually. These four options are all distinct and (by definition) complete.

When an intellect comprehends the form of a squirrel, it remains an intellect: there is no change in substance. That there exist material objects which manifest particular forms by change of substance is obvious. That there are immaterial objects which can manifest particular forms without change of substance is harder to prove; and it may be questioned whether such things exist. The definition, of course, remains valid whether intellects exist or not. That they exist does not, however, require philosophical demonstration; the very fact that we are capable of abstract thought indicates that we possess intellect; and if we possess an intellect then intellects can exist.

Now form is an example of an abstract idea. It thus follows that everything that is capable of intellectually grasping form is capable of grasping at least some abstract ideas. Does it follow that everything that is capable of grasping other types of abstraction, such as

arithmetic, is also capable of grasping form? It would seem so, because knowledge is first of all derived from our senses, and only once we have some kernel of knowledge from our sensual data can we increase it through abstract reasoning. Yet form is the first place where sense encounters intellect; to understand what we see means grasping the form of what we see. It is only after we can comprehend multiples of the same type of object that we can consider number, and from there to the more abstract concepts such as fractions, complex numbers, and negative numbers, and from there to the higher forms of thought such as mathematics and logic. It seems to me that this must be true for every intelligent material agent, not just mankind, for the ability to engage in higher reasoning is worthless if it cannot engage with the surrounding would which it needs to sustain it; and truth in reasoning requires a true premise which must be derived without reasoning and can therefore only come from observation. □

DEFINITION 4.4.3. **Will:**
A will is related to a tendency towards an end. In a material substance, this is related to the final cause: certain material substances have the possibility to change into certain other types of substance or certain other potentia within the same substance. A will, however, is more associated with immaterial substances; it refers to any tendency towards an end which is not a final cause in matter. It may be defined as the element of some immaterial form that has the ability to institute change in other beings. We may define several types of will:

- **External will:** An external will is where changes of a being's state are directed by something wholly external to that being. A gas in a room, for example, is directed towards equilibrium by an external will (described by the laws of thermodynamics).
- **Internal will:** An internal will is where the changes of a being's state are directed by something wholly or in part internal to that being (though the direction may be mediated through physical law). For example the will of an animal is internal. 'Internal' here refers to the source of the will, not necessarily its target or effects.
- **Determined will:** A determined will is a will where the end result of the change of the object is deterministic; i.e. given a particular initial state, the will is directed and can only be directed towards a solitary final state for all initial states. There is a fixed sequence $A \to B \to C \to \ldots$.
- **Undetermined will:** An undetermined will may be directed to more than one final state (with a significant probability for each decision) given for each initial state. Circumstance A implies B_1 a certain percentage of the time, B_2 another percentage of the time, and so on, and there is no way even given full knowledge of the situation to know which effect the undetermined will will select; the best one can do is express the possible choices probabilistically.
- **Intellectual will:** An intellectual will is where the will is directed to seek desires on the basis of what is grasped by the intellect.
- **Free will:** A free will is intellectual, undetermined and internal; where the ultimate final state of the being is selected only on account of factors internal to that being, and it it is not possible to predict the actions of someone with free will through calculation and even a complete knowledge of their state immediately prior to that decision.

△

Objection:
We have definite knowledge of choices that were made in the past. Those choices were

therefore determined. Since free will is inconsistent with determined choices, free will is impossible. •

Response:
Free will is about the inability to deduce what future events will be based on complete knowledge of the present and of physics. It states that we cannot predict the outcomes of a choice made freely. But knowledge can be obtained through other means than prediction. If our knowledge of the choice came through some other means, there is no contradiction between the results of the choice being both unpredictable and certainly known. Given complete knowledge of Marcus Brutus and the other conspirators in early March 44BC, it is impossible, based on that knowledge, to predict that they would have stabbed Caesar on the Ides of March, since they had free will and could have decided either way. However, we know that they did make one particular decision; and we know what that decision was, not through prediction, but through our awareness of past events. Free will solely implies an inability to predict the future of agents with free will; but if knowledge of the future can arise though other means than prediction, then future knowledge is not incompatible with free will. There is thus no contradiction between saying that Brutus choose freely and that we know which action he took. ○

DEFINITION 4.4.4. **Indeterminacy:**
Indeterminacy means that the end state is not inevitable for each possible initial state; i.e. for a being in an unstable state $|I_1\rangle$ there is non-zero probability that it may end in each of the states $|F_1\rangle, |F_2\rangle, |F_3\rangle, \ldots, |F_n\rangle$ for $n \geq 2$ with $\langle F_i|F_j\rangle = 0$ for $i \neq j$ (this last equation means that F_1 and F_2 are orthogonal states, i.e. they are not just the same thing written in different ways). In short, for an indeterminate system, given complete knowledge of the system at one moment in the sequence of causality, it is not possible to predict with certainty the state of the system at the next moment in the sequence of causality. A determinate system is where one can make such a prediction. △

DEFINITION 4.4.5. **Freedom:**
Freedom may be defined as the ability to do whatever is desired by the will. An unfree being is one whose desires are constrained by some limitation or external factor. The word desire may be interpreted here in one of two senses, either in accordance with the will of the being (the modern usage), or in accordance with the final cause of the being (the ancient usage; final cause will be defined in definition 4.5.10). △

DEFINITION 4.4.6. **Perception:**
Perception is the ability to observe objects external to oneself through the senses. For us, this includes the formal abilities associated with sight, hearing, touch, taste and smell. An object that is rightly perceived by the senses (that is when the senses are functioning correctly: for example hallucinations are not a false perception but a defect in the senses; illusions are are true perception that is misjudged by the self) is accurately transferred from sense into imagination. △

DEFINITION 4.4.7. **Imagination**:
Imagination is the ability to visualise something within oneself. This may correspond to something that exists in reality, or it may be something that existed in the past, or something that will exist in the future, or which will never exist. Δ

DEFINITION 4.4.8. **Self-awareness**:
Self-awareness is the recognition that we exist, knowledge of our past and the awareness of the will, intellect, imagination and so on. It is the focus of unity within the mind. Δ

Discussion:
One could, perhaps add memory to the list of components of the mind; this is, however, a matter of definition, and I have included memory within self-awareness (for those beings where memory is relevant; for example we only require memory because we perceive each moment in time in succession, and our only access to the past is through memory and to the future through prediction. However, a being outside time, one which perceives moments together rather than in succession, would have no need for memory because it would always have direct access to both the past and the future).

The first question is where thoughts or ideas enter into this pattern. I will define thoughts as the means by which we grasp things external to the mind. That is to say, the object of our thoughts is not part of ourselves; we may think about dogs, or calculus, or the pain in our big toe, but none of these are part of our mind. Even when we think about our mind, we do not consider it *per se*, but only of an abstraction that represents our mind. This is in contrast to the view set forward by Locke, and continued by his successors such as Hume,

> I must here in the entrance beg pardon of my reader for the frequent use of the word *idea*, which he will find in the following treatise. It being the term, which, I think serves best to stand for whatever is the object of the understanding when a man thinks. ... I presume that it will be easily granted me, that there are such ideas in men's minds: every one is conscious of them in himself; and men's words and actions will satisfy him that they are in others[13].

The difference between Locke's position and that which I am advocating (my position is, of course, by no means original to me) is subtle, but important. For Locke, ideas, which reside within us, are the object of our thought. For me, ideas, which reside within us, are the means by which we think of objects external to us. From Locke's position, we see that we do not directly perceive the outside world. We perceive ideas within us, which may or may not refer to the outside world. Colour, shape, texture and so on are thus properties we give to objects but are not part of the object themselves. For me, these are properties of the object we perceive; for example colour refers to the dominant wavelength of light which is reflected off the surface of the matter, which in turn comes down to the atomic structure of the material: it is a property of the matter. For many of those who follow Locke, it is something which we add to our idea of the matter, but has no part in matter itself. Equally, Locke's position generates a distinction between the world of our ideas and the world of reality (if there is such a thing as a world of reality external to ourselves); a distinction which is absent in my view of the mind. Locke's definition of *idea* then leads to the question of what relationship there is between the world of ideas, and the world of reality. Numerous solutions have been proposed to answer this question; I leave it to those better versed in the intricacies of modern philosophy than I to judge how successful these

[13]Locke, *An essay concerning Human understanding*, Introduction.

attempts have been. But this question is entirely absent in the understanding of the mind which I am advocating here; it simply does not occur.

It is traditional to divide those things which are the objects of our ideas and thoughts into things sensible (material quantities discernible to our perception) and things intelligible (abstract quantities discernible to the intellect). Of course, there may well be relationships between these two categories; for example we can both see triangular objects and consider triangles in the abstract; on the other hand we cannot perceive algebra but we can think of it abstractly. This division of objects naturally leads to a distinction between our perception and intellect. This view has been challenged by modern philosophers, who generally fall into two camps both opposed to it. The first camp contains people such as Hume, Berkeley and so on who deny the distinction within our mind by denying that we have an intellect (as I have defined it here). In short, we are not able to grasp abstract quantities: what we grasp are particular examples. For example, when we think of triangles, we imagine (and, for the purpose of this discussion, I shall treat imagination as part of perception, for the obvious reason that if Locke's view of ideas discussed in the previous paragraph is correct, as Hume assumes, there is no fundamental difference to our awareness between imagination and perception) one particular triangle, not triangularity in the abstract. This is, of course, false, as I am perfectly capable of thinking about triangularity without visualising any triangles (a triangle can, for example, be represented by three pairs of numbers, and it is easier to prove certain theorems regarding triangles in this way) or by visualising several triangles at once. What links these together is an abstract quantity, derived from the forms of triangular objects, and it is difficult to accept why we relate numerous different objects as triangles unless there is something shared between them. This something is not a matter of one particular object, but relates to an infinite number of objects, or would still exist even if there were no triangles in nature or it was impossible to imagine them. For example, I can think about objects analogous to triangles within twelve dimensional space (that is, objects whose corners are defined by thirteen sets of coordinates, just as the corners of two dimensional triangles are described by three sets of coordinates) even though I cannot visualise such objects and even less perceive them. Nonetheless, Hume denies us the ability to think abstractly:

> Let a man try to conceive a triangle in general, which is neither isosceles nor scalene nor has any particular length or proportion of sides; and he will soon perceive the absurdity of all the scholastic notions with regard to abstraction and general ideas[14].

Hume's problem is that he confuses the intellect with the imagination. This mistake is at the heart of many of the errors which Hume made (and one of the reasons why his philosophy cannot be accepted), such as his views on causality. When we imagine a triangle, we think of one particular triangle. But it is perfectly possible, and mathematicians do it all the time, to reduce a triangle to its abstract properties, and manipulate those concepts and draw relationships between them symbolically, without necessarily imagining any triangle. Thus the intellect can grasp the abstraction of triangularity in general, even if Hume is right in saying that the imagination cannot.

Of course, Hume would not just apply this to triangles but all types of object. Now, of course, I can think of a triangle as an object whose corners in a Cartesian Euclidean geometry at (x_1, y_1), (x_2, y_2) and (x_3, y_3) are connected by straight lines, and perform all manner of calculations concerning it and discern all manner of rules about it and grasp all manner of possibilities about it before setting the coordinates to some particular values, regardless of whether this triangle may in practice be isosceles or scalene. I respond to objections that this is not a triangle at all but some mathematical abstraction by observing that that somewhat proves my point. But nonetheless, Hume concluded that if there is no place for abstract ideas, then everything depends on our perception (although Hume

[14]Hume, *Enquiry concerning Human Understanding.*

excludes mathematics from this). Only that which we can perceive can be established as true. This is the doctrine of empiricism.

> When we run over our libraries, persuaded of these principles, what havoc must we make? If we take in hand any volume; of divinity or school metaphysics, for instance, let us ask *Does it contain any abstract reasoning concerning quantity or number?* No. *Does it contain any experimental reasoning concerning matters of fact and existence?* No. Commit it then to the flames: for it can contain nothing but sophistry and illusion[15].

This principle thus, if true, destroys all metaphysics (including materialism and atheism), ethics, and so on. I also believe that it undermines theoretical physics. Experimental physics observes regularities in the universe; it deals with matters of fact and existence; fine. But as soon as we try to draw patterns and inferences from the data; to predict what would happen even without conducting an experiment or in distant corners of the universe which we can't observe, then all such works, if Hume is taken to his logical extreme, should be committed to the flames, for they do not deal with real objects but abstractions. Thus, if he took his principles to their logical conclusion, Hume would cast all of Einstein's and Dirac's and Feynman's works into the flames, for they are not solely idle mathematical curiosities but are intended to be applied in the real world, and not experimental reasoning. The mathematics we deal with goes far beyond quantity and number; we introduce abstract concepts such as gluons, metrics and virtual particles which have never been directly observed and are in principle unobservable. The laws of physics are not directly observable (only their effects when applied to matter: we do not see a differential equation such as the Schrödinger equation, we see the outcomes of various experiments; outcomes whose distribution can be predicted from the Schrödinger equation or more advanced quantum theory). The Laws of Physics would just be a human construct rather than something embedded within the structure of the universe. This, of course, leaves open the question of why they are so successful.

Hume has, however, made two mistakes. The first is to fail to properly distinguish certainty, and different degrees of reasonable doubt and unreasonable doubt: there is much use in speculation which goes beyond the observed evidence but tries to explain it. The second, and more important, mistake is to dismiss abstract reasoning entirely, and not see that a combination of geometrical and empirical reasoning can lead to places that either alone could not reach. There is thus a second test of truth beyond experiment. Is this reasoning self consistent? Are the definitions reasonable? Are the premises or conclusions sound? If all of these are satisfied, or it is possible that they may be satisfied, then I do not think we are right to dismiss such work as valueless. I certainly do not wish to demote the importance of experiment, as I will discuss later. I am only saying that experimental reasoning and mathematical reasoning is not everything. (As an aside, I would comment that *Can the premises behind Hume's empiricism be established from abstract reasoning concerning quantity or number?* No. *Are they based on experimental reasoning concerning matters of fact or existence?* No. It is very fortunate for Hume's disciples that I do not approve of burning books; Hume's works should be preserved, albeit moved to the comedy section of the library, so frequent and ludicrous were his mistakes.)

Hume's arguments, of course lead to the general position of nominalism: that there are no general truths only particular truths; and positivism: that only the experimental sciences can teach us anything. Both of these, as I have explained, are thoroughly incompatible with my general thesis, and if, at the end, the reader thinks that there is some merit in this work then they must, along with the premises of Hume and Locke which led to them, be rejected.

But before I discuss where experiment fits into my scheme, let me briefly mention the opposite error from Hume. This is the position that there is an intellect as well as perception, but the two are wholly separate. The world of ideas is distinct from the world

[15]Hume, *Enquiry concerning Human Understanding.*

of reality. There are truths known inherently to the intellect, and not just mere tautologies, without requiring any sensual input. Various forms of this doctrine, *dualism*, we held by the likes of Descartes, Kant, and, before them, Plato. These positions lead to the famous problem of how the intellect and perception are related to each other. I am, of course, advocating that both perception and intellect exist, and that are distinct from each other. I do not, however, agree with this dualism, because I believe that they are distinct but not fundamentally different or separate. Rather, their union makes a mind.

I said before that an abstract argument should be accepted if the premises are sound, and considered distinctly plausible if the status of the premises are unknown but the conclusions are sound. What do I mean by sound? This may be a definition, or tautology, or based on experimental observation. Equally, the conclusion may be accepted if it is in agreement with the self-evident or experiment.

In short, almost everything abstract we ponder with the intellect is spawned from something else: this may either be another intellectual thought, or it may be from something we perceive with our senses or imagination. We see particular objects; from that we learn general principles; from those we can start to engage in abstract reasoning. It is because I have seen that I can imagine; because I can imagine that I can extend my imagination to things as yet unseen; because I can extend that I can discern patterns and think abstractly. It is because I have seen a triangle that I can think about triangles in the abstract; because I have seen a definite coordinate system as a grid in the external world that I can visualise an abstract coordinate system, and extend it beyond Cartesian coordinates to any system I fancy; because I have seen two or three dimensions that I have learned how to treat them mathematically, and because I can treat two or three dimensions mathematically that I can extend those techniques to four, eleven or twenty-six and a half dimensions. In short, the distinction between a perceptual thought and an intellectual thought is one of relationship. The perception comes first; the intellect then builds on it. True, I may then use the intellect to construct an image which I then create and perceive (for example, if I were to construct a work of art – and sufficiently skilful, which I am not, that what I constructed bore some vague similarity with what I had previously imagined); but this is a different process, creation rather than understanding. In understanding, the order is perceive – imagine – abstract thought which can be extended well beyond what it perceived and even what is possible to perceive (but still, nonetheless, true); in creation it is abstract thought – imagine – perceive. In each case, the principal differences are not just our mental processes (which are attuned to the type of thought) but the order of the thoughts and thus the relationship between them.

Thus, although the different qualities of mind listed above are unified into one substance, they are distinct both in their object and their relationship to each other. In short, Hume identified the mind with perception. Descartes divided the mind into intellect and perception. There is, however, another possibility, and that is what I am adopting here.

It should also be obvious that while I consider metaphysics to be useful and important, I do not believe that it should be a free-for-all where anything goes. Just as perception is related to the intellect, so physics and metaphysics are related. If a metaphysical system cannot reach physics in either its premises or conclusions (or, worse, that it reaches a wrong physics), then, barring the limited good that can be achieved from self-evident truths, it is useless. Physics explains what the laws which govern the material world are. Metaphysics must either start from those laws (or some other, non-scientific, observation) or finish by explaining why those laws are as they are. If not, it is at best speculation and at worst dangerously wrong. Every metaphysicist must also therefore first be knowledgeable about the most modern established understanding of physics. □

DEFINITION 4.4.9. **Rational agent**:
A rational agent is a being which possess essentially a mind and whose motions are wholly or in part or in principle controlled by that mind.

By 'in principle' I mean that it is not foreign to the nature of that type of being. For example a man in a coma does not consciously control his body, but remains a rational agent because he is still a man, even if disordered, and it is of the essence of man to be a rational agent. Even if he is not in control of his body at that present moment, he will still be able to at other times, or, more precisely, if some other potentia of his form were actualised. Δ

DEFINITION 4.4.10. **Simple**:
A simple substance is one where it cannot be dissected into individual components which are beings in themselves and differ in substance from the original. Δ

DEFINITION 4.4.11. **Composite**:
A composite substance is one which can be dissected into individual components which are beings in themselves and differ in substance from the original. Δ

DEFINITION 4.4.12. **Life**:
A living organism is a single inhomogeneous substance which has the tendency to change its accidents itself (although it may also be changed by external factors) without any change to its essence; a non-living being can only be changed by factors external to it. Living organisms have the tendency to grow, ingest matter, reproduce, and sustain their substance against damage or invasion, offering resistance, at least for a limited time, against thermodynamic decay. Δ

Discussion:
A crystal is not a living organism because although it may grow it is a homogeneous rather than inhomogeneous substance, in that it is possible to divide the crystal into portions which have the same structure as each other.

A particle which undergoes spontaneous decay is not a living organism by this definition because the substance is not left unchanged during the decay.

A corpse is not a living organism because it is not self-sustaining.

The liver is not a living organism because it is not self-sustaining by itself unless in conjunction with the rest of the body; it is part of a living substance, but not a living substance in itself.

Note that the tendencies of growth, metabolism and so on are directed towards changing the being itself, from one potentia of the substance to another, but do not directly point to an effect outside itself. The tree has the natural tendency to grow itself, but not to cause something else to grow. This is distinct from non-living substances, which also have inherent tendencies to particular ends (as a radioactive particle has the tendency to emit an alpha-particle), but these ends are always directed to an effect outside the substance (so when the particle decays, it emits a helium atom, something external to itself). There is thus a distinction between tendencies towards internal and external effects (which are known as *immanent* and *transient* causation respectively). The traditional definition of a living organism is that only life displays imminent causation, while all things (both living

and non-living substances) display transient causation. □

DEFINITION 4.4.13. **Soul**:
The soul is the form of a living organism; i.e. a form which contains (but is not limited to) the description of the properties and various arrangements of matter which distinguishes life from non-living beings. Each individual living organism has its own individual form, and therefore its own individual soul. △

Discussion:
Obviously, there are others who define the soul differently – as some sort of vital energy or *qi* which gives life to the body. Defining a soul in terms of such a vital energy or principle immediately falls into problems when confronted with biochemistry. Life is known to exist on account of the interactions and motions of the various molecules which make up the body, processes which are now well understood without any reference to *qi*. The idea of *qi* therefore can only be correct if it is an alternative way of saying exactly the same thing as biochemistry – the same idea expressed in different words; or at least a good approximation to or simplification of it. This can be determined by whether these make the same predictions for how the body reacts in certain circumstances and start from consistent premises. Whether this is so is a question for those who are experts in both some vital energy model and the biochemistry of the human body, and I am not, on either account. But, to my knowledge, no model of the human body based on vitalism has survived this sort of scrutiny.

But this is not what I am suggesting here. This definition of the soul, which I have adapted from the best of the philosophical theologians, is very different. A living body contains the same matter as a dead corpse; at least until the effects of decomposition become significant. It has the same number of electrons, protons, neutrons and so on as the corpse. What is it then that distinguishes the living body from the corpse? At the risk of being a little imprecise, the arrangement of the matter: their locomotion, position, energy levels, molecular structure and so on. The complete set of possible arrangements of matter that allows the electrons, protons and neutrons to form a living man is the soul.

Of course, biologists know what the difference between a corpse and a living body is, if not fully then at least to a great deal of accuracy (that the process of death is not instantaneous across the body does not change this point: my argument does not require it to be instantaneous, but only that there is a clear distinction between its start and its end). Does this undermine the concept of the soul? No, it supports it. The question of 'What is a human soul?' or 'What is this particular man's soul?' is one for metaphysicists to ask (since they provide the definitions) and for biologists to answer, either completely or partially depending on whether consciousness can be explained fully by neuroscience (a debate about which I know little and therefore don't want to take sides). And, for the most part, this question has been answered, even if not expressed in these terms. In short, the soul as understood in some major streams of classical theism is something that both can be and largely has been constructed from biological measurement – the biologists did not realise that that is what they were doing because of a conflict of definition. Only if the biologists were to find that there is no microscopic difference in the arrangement of matter between a living body and a corpse would the concept of the soul and distinction between soul and form be undermined.

Of course, this leads to the question of why this particular soul gives a living being while that particular form of the same matter does not. But that question is in no way undermined by the concept of the soul. The soul is the formal cause of the man; why that particular arrangement represents life is a matter of the efficient and final causes. Again, it

is a question to be posed by metaphysics and answered by biology; and the answer, whatever it is, cannot undermine the definitions that led to the question.

But one further comment needs to be made and emphasised: the soul is immaterial. It impresses itself on matter, but it is not in itself made of matter. Just as one can model an hydrogen atom with the equations of quantum mechanics without needing to observe a hydrogen atom; so one can discuss the form of a living body in the abstract without reference to that living body (even though a living organism is far more complex than an atom, the same procedure of writing down an equation which describes the possible wavefunctions of all its constituent parts and solving it can in principle be done, even if none of us are smart enough to actually do it; and of course there are other, and perhaps better, ways of considering a soul in abstract thought). It is also possible for the soul to survive death; if it is just a description of the possible arrangements of matter, then this can in principle be preserved (for example in principle, though not in practice, in a computer memory bank; or alternatively in a mind). Equally, our memories and concepts of who we are are preserved in our soul. □

DEFINITION 4.4.14. **Animal soul**:
An animal soul is the form of an animal, distinguishing it from plants, fungi, bacteria and other living organisms. Animals are constructed of cells contained within a membrane and multi-cellular, which separates them from bacteria and viruses, and they lack a rigid cell wall. They are capable of digesting food in a specialised organ, and do not photosynthesise. They are also capable of self-locomotion, and many fulfil their tendency towards reproduction sexually. Most animals will contain a nervous system, sensory perception, and are capable of responding to stimuli. They also have an instinctual desire to satisfy their appetites for food and drink (among others), and are capable of ordering their actions to fulfil these appetites. △

DEFINITION 4.4.15. **Rational soul**:
A rational soul is a soul which contains a mind. △

DEFINITION 4.4.16. **Rational animal**:
A rational animal is a substance whose form is a soul, rational, and animal. △

DEFINITION 4.4.17. **Spirit**:
The spirit refers to the immaterial aspects of the mind. △

Discussion:
That the mind has immaterial aspects as well as material aspects (such as sensory perception) is clear, since the intellect understands by taking on the form of what is understood, without taking on the matter, and the form is immaterial. It does not matter whether this immaterial aspect is derived from the (material) impulses of neurons or the material aspects of the brain are driven by the immaterial, or a combination of the two. There is no contradiction in saying that the material can give rise to the immaterial, as the arrangement

of matter always determines an immaterial form; equally we could say that the immaterial form imprints itself on prime matter to give the material object.

Also note that since the mind contains an intellect, that is the capacity for abstract thought, and as far as we know we are the only species of animal capable of abstract thought, mankind is the only rational animal. Of course, that is just based on our current knowledge. It may be that there are alien life forms that are also rational animals; it may be that some other animals on this planet possess an intellect in addition to the basic sense perception, memory and their ability to form correlations between different material events, mimic other beings, home in on prey or flee predators, or recognize various geographical or other features. Many animals certainly have enough intelligence to participate in their surroundings, recognize and react to circumstances and plan ahead; but as far as we know that intelligence does not extend to the capacity to extrapolate from the perception of triangles to the concept of triangularity, or from counting rabbits to the concept of irrational numbers, or from that to algebra or calculus, or even artistic representation. It is this capacity for abstract reasoning (a capacity not always realized – as shown by the existence of certain politicians) which differentiates a rational animal from a mere animal. □

4.5. Causality

DEFINITION 4.5.1. **Existence:**
The primary sense of existence is: a being can be said to exist if it is capable of interacting with another being, either modifying the state of the other being, or having its own state modified by something else. In this sense we say that a table exists, or a dog exists, or a ball exists.

A second, looser, use of existence describes a relationship or interaction between two beings which exist in the primary sense, for example we might say in a secondary sense that a friendship, or a gravitational force exists.

A third sense of existence refers to abstract concepts, and is wholly separate from any beings which exist in the primary sense; for example the word may be used in an algebraic relation, or to describe the form of a material object. Here it mainly denotes the logical consistency of an abstract ideal.

These senses of the word are distinct, and in my main argument I shall chiefly use the first sense of the word.

Note that for almost all substances, existence is an additional accidental property of the being. The form of a being exists (in the third sense) regardless of whether the being exists in the first sense; the concept of a dragon does not depend on the existence of any dragons. Equally, beings which once existed may later cease to exist; for example Socrates existed in the past, but no longer exists. Thus Socrates once had the property of existence, but no longer does so. A being which exists essentially rather than exists accidentally could not cease to exist or come into existence, as this would mean that there would be times when it would not be in conformity to its essence; equally it must exist over all space. As all material beings are local in space (from observation) and time, they can only exist accidentally rather than essentially. △

Objection:
As Kant showed, existence is not a predicate or property. Every predicate adds something to the concept of a being, for example 'This post-box is red' adds the concept of redness to the idea of the post-box. However, existence adds nothing, because in the statement

'This post-box exists', the existence is already implied when we identify the post-box. Thus existence adds nothing to the concept, and is not a predicate or property. •

Response:
The argument here fails to distinguish between *essential* and *accidental* properties, but assumes that there are only *essential* properties. It also assumes that it is properties, not form, which determine the nature of the being. These assumptions are false, and thus the argument is invalid.

The form of a material object can be understood independently of the object's existence; for example we can talk about the form of a unicorn. To say that the unicorn exists is to say that that form is actualised. Equally, we can still say that a non-existent post-box is red, because by identifying the post box, we consider its form, and we can compute the necessary properties of that particular post-box, such as its redness, from its form. However, existence is an accidental property of the post-box. Likewise 'reclining' is an accidental property of a man. We cannot say from its form whether or not the post-box exists, or whether a man reclines. Thus existence, or reclining, tell us something about the actual state of the post-box or man, and are useful expressions. Even for essential properties, such as the post-box's redness (essential for this particular type of postbox), the statement 'This post-box is red' is ultimately a tautology, but still useful because it is not self-evident just from knowledge of the form, but has to be computed, e.g. from the energy band structure of the paint on the surface of the post-box.

For example, I might be shown detailed plans and images of the Taj Mahal, and come to a complete understanding and appreciation of it. However, I cannot know from just those plans whether there is truly such a building in India, or if it is just a fantastical concept. I can understand the form and study its history without knowing whether or not the building exists; to be told that it does, in fact, exist will surely add to my appreciation of it.

I will give a fuller response to this objection in section 12.4. ○

DEFINITION 4.5.2. **Potentiality**:
The potentia are the possible states of being of a particular substance. They describe what changes are possible while still remaining the same substance.

Potentiality describes the states that a being could be; so a red ball is potentially a white ball if it is painted, but not potentially a bunch of flowers. An ensemble of particles with total spin $\frac{1}{2}$ is potentially an ensemble of those particles with a different arrangement of the matter and again total spin $\frac{1}{2}$, but cannot become an ensemble of particles with total spin 1 (unless another particle with half integer spin is added to the ensemble, but this would effectively mean that we were discussing the larger ensemble, so a different system). The different arrangement of matter is another potentia of the system, the new ensemble is not a potentia of the system because it involves a discontinuous change in the system, in this case either adding or taking away a particle. △

DEFINITION 4.5.3. **Virtual existence**:
An object can be said to exist virtually within another substance if it is a final cause of one of the potentia of that substance but not itself another potentia. Suppose we have being A in a potentia A_1 which either decays into being B or emits being B while changing to a potentia A_2, where B is not of the same substance as A. In either case, the decay to something that exists virtually involves a change in A_1, either its destruction or a change to a different potentia. If B exists within A without causing a change in A, we say that it

exists intellectually within A (definition 4.4.2). Δ

DEFINITION 4.5.4. **Actuality**:
Classically, actuality describes a being in the state that it is; of all the potentia, it describes the one which actually is at a given moment in time.

Only a being that exists can be actual; only something which is actual can exist. Δ

DEFINITION 4.5.5. **Motion**:
Motion is movement between one potentia and another. At one moment in time, the potentia X_1 was actual, and X_2 was only a possible but not actual state of the being. At the next moment of time, potentia X_2 becomes actual. Motion is that change from state X_1 to state X_2.

For example, X_1 and X_2 might represent different heights above the ground. If we have something falling off a table, then at the first moment in time it will be at the top of the table (X_1), then in the next moment of time gravity will do its work and the ball will be at a slightly lower location, X_2. This type of motion is known as local motion or locomotion, movement from place to place.

One of the problems we face in understanding Aristotle's metaphysics is that in the scientific revolution of the sixteenth and seventeenth centuries, it was assumed that locomotion was the only type of motion. Thus we have come to confuse the two words, and treat them as synonymous. However, the Aristotelian concept is much broader than this, and it is therefore necessary to distinguish the concepts of motion and locomotion. We can think, for example, of a banana ripening on a shelf. First of all it will be green, then yellow, then gradually becomes brown. X_1 is the state of the banana when we first see it, when it is green. X_2 is the state in the next moment of time, when it is slightly less green, and a bit more yellow. And so on. Of course, you might regard this example as being a bit silly: the state of the banana ultimately relies on the states of the various different molecules that make it up, and what is happening to them is far more complicated than a simple change of colour, and (for all we know at this point in the discussion) might ultimately be reducible in practice to locomotion of the constituent parts. But, of course, we don't know that's true either, until we begin a detailed chemical analysis of banana skins. The point I am trying to illustrate is that Aristotle's philosophy allows for different types of motion than just locomotion. Whether that is useful or not is something we will have to figure out when we know enough physics. That is much more physics than was known by Aristotle, but also more than was known by the early modern scientists who first proposed that locomotion was the only type of motion.

Aristotle allows three different types of motion:

(1) Remaining in the same state. Technically, this isn't motion at all, but needs to be listed here for completeness.
(2) Artificial motion. This is when the being interacts with something external to it. For example, when I lift a book from the floor to the table, I am applying artificial motion. I'm the one doing the work.
(3) Natural motion. This is motion that arises from the internal nature of the being, and is related to the being's final causes. When the book falls off the table, it undergoes natural motion downward, since among the final causes of the book is to rest as close as it can to the centre of the universe (at the centre of the earth – please bear with me while I express things as Aristotle did in his physics even though it is completely outdated: I am, after all, trying to explain his philosophy,

so I will use his example. But bear in mind that the philosophical concept is more general than this example). This motion is not caused by anything external to the book, but arises from the nature of the book itself.

△

DEFINITION 4.5.6. **Event:**
An event is that moment when a potential is actualised, or a previously actualised potential ceases to be actual (usually at least one instance of both of these occur at the same time). It corresponds to a change in something. △

There are four primary types of cause: efficient, final, material and formal. We will also discuss instrumental causes, but these can be expressed in terms of a sequence of efficient causes. Every change requires a cause; every material being in every state has each of the four causes.

DEFINITION 4.5.7. **Efficient cause:**
Efficient causes describe the substances and the state of that substance that precede a being's existence. For example, the efficient cause of an anti-neutrino may be neutron that underwent a weak nuclear (β) decay into a proton, electron, and anti-neutrino. △

DEFINITION 4.5.8. **Material cause:**
The material cause is the matter out of which the object is constructed. △

DEFINITION 4.5.9. **Formal cause:**
The formal cause is the form of the being. △

DEFINITION 4.5.10. **Final cause:**
There are, unfortunately, five distinct senses in which the word 'final cause' is used. Confusion between these senses has led to numerous errors, so it is worth spelling them each out in turn.

(1) A list of the results of the possible changes that a substance can undergo. The most important members of this list are the internal tendencies of matter; those changes of state which occur without being directed by something outside the being. For example, an apple tree has the tendency to produce apples; an apple tree in a state which bears fruit is a final cause of an apple tree in a state where it is in blossom. The possible decays or excitations of a relatively simple particle can be identified with the final causes of those particles. There is a change of state: either the particle itself ceases and becomes other particles, or it remains but new particles come into being, or the particle moves from one actuality to another. This represents a state towards which a single being has a natural tendency.
(2) The thermodynamic equilibrium of a complex system of particles. This represents a set of states towards which a system of many beings has a natural tendency.

(3) The function of an object (the final cause of a pen in this sense is to make marks – such as writing – on a paper, as that is what it was designed to do, and one of its effects when used properly).
(4) The purpose or goal of a being according to a mind (for example the mind or minds that created it or use it or set it into the state it is in, or an earlier state, or direct its motion).
(5) We may discuss the will as an example of final causality: if the will desires that a ball hit a stump, then the related final cause is a state where the stump has been dislodged through contact with the ball. As before, the will need not be directed towards only one effect. We should again think of this sense of the final cause as a list of states, but this time these states arise as a result of mental processes. (I haven't discussed whether mental processes are always reducible to physical processes. If they are, then the will would ultimately be reducible to the first sense of final causality. But it is would still be a useful distinction, as there is a difference between knowing the microscopic details and seeing the macroscopic picture. Of course, if mental processes need not be reducible to physical processes, then this sense of final causality is only analogous to the first two, but it would play a similar role.)

The first of these is an opposite of the efficient cause; a final cause of the neutron is the β decay into the proton, electron, and anti-neutrino; while the efficient cause of the electron is the neutron. A final cause of two particles moving with particular momentum may be the particles moving with different momentum subject to the conservation of energy and momentum. Note that just as a particle may decay in several different ways, it may have numerous final causes; however each individual particle only has one efficient cause. There are numerous different states something could change into; but in practice it came from just one particular being or conjunction of beings. Thus final and efficient causality are distinct, and not simply different expressions of the same thing.

The second sense of final causality arises from the combination of the final causes of the objects in the first sense continued to a large time (so the final causes of the beings at a particular time give us the system as it is at the next time interval; we then take the final cause of that system, and continue until we which equilibrium). In this sense, the final cause of a gas in a room is to be spread out with equal (within certain minuscule fluctuations) pressure and density of particles within each sub-region; the final cause of the arrow in flight is to be at rest in its target. The first sense of final causality refers to single particles over short time periods; the second sense applies to systems of many particles over long time periods, but the two are closely related, in that we can derive the final causes (in the second sense) of a system of particles from knowledge of the final causes (in the first sense) of each individual particle in that system.

The fifth sense of final cause appears to be very different; it is related to the will; indeed the will is to an immaterial being what the final cause in the first sense is to a material being. Clearly, were there no minds, then this definition would not make sense, but given that there are minds we must also consider it.

For example, the final cause (in the fourth sense) of cooking is to prepare food; the final cause of food (in the third sense) is to be eaten; a final cause (in the second sense) of food being eaten and digested is to nourish the body, and it achieves this through many interactions of fundamental particles according to their final causes (in the first sense). Note that it does not matter whether the food is in practice eaten or discarded; the purpose for its preparation remains the same, that it should be eaten. (Equally, if garbage is eaten, that does not modify the purpose for which it was made; to be sent to the dump or recycled.) This fourth type of final cause depends on the desires and purposes of a mind, and therefore on the will. The fourth sense of final cause is thus related to the last. The third sense of the term is distinct as well, and lies as something of a bridge between the others. For a

complex being, we may ascribe functions to its various parts, for example the function of a heart within a living being is to pump blood. This can be seen firstly by studying its final cause of its components in the first sense; when correctly stimulated, the muscles of the heart contract in such a way that causes the blood to circulate the body; and then to note that the living body would fairly quickly cease to be a living body if the heart ceased producing this effect. This type of final cause thus can be derived from the form of the entire substance (the living organism) and the final cause (in the second sense) of the part (that the heart moves the blood in this way when it beats).

We should therefore distinguish the first three senses of final cause, which do not directly depend on the presence of a rational agent, from the fourth and fifth, which do. From this point, when discussing physics, I will use the phrase *final cause* or *teleology* just to denote one of the first two senses of final causality listed above (which one I am referring to should be obvious from the context). The third sense I will call a *natural purpose* and the fourth an *intellectual purpose* or *design*. With regards to the fifth sense, I will just here note the similarity between the will and the final cause. When discussing the final cause of an immaterial being (I haven't yet discussed whether there are such beings, so at the moment this comment is purely hypothetical), then I will mean the fifth sense, since the first two cannot apply. The first, second and fifth senses each imply a tendency towards a set of ends, but are manifested in different ways because of the different natures of the substance.

Note that a being may have more than one final cause (even in the same sense of final cause; for example a particle might have several different possible decay channels). Equally note that although a being may have a final cause, it does not necessarily mean that that will happen. That a down quark has the power to decay into an up up-quark, anti-neutrino and positron does not mean that it will in fact do so. Final causes give possibilities or even probabilities, not certainties[16].

It is important to emphasise the difference between the first and fourth senses of final causality. There is a clear distinction between the concepts 1) Causes tend to produce certain effects and not others; and 2) Causes tend to produce certain effects and not others because of some natural purpose, intellectual purpose or perhaps design. The first question is a matter for physics to answer, and in my discussion of physics I will only address this issue. The second question is a matter of theology or natural philosophy, and I shall defer it to after the discussion of the first question.

One of the many errors made in the early modern period was to confuse these two senses of final causality. Partly this was because some classical writers used the phrase *final cause* to describe the first of these meanings (I would place Aristotle and his disciples in this category) and others used the phrase to denote purpose (I would put the Stoics and Cicero in this category). The two are, however, very distinct. In the early modern period, the forerunners of classical science discussed what place final causality had in natural philosophy (what we now call science); this was a debate that extended beyond the time of Newton, though the most important contributions to it were made before Newton. One side of this debate argued that the laws of physics proceeded without any reference to purpose; and therefore final causality was at best superfluous. The other side of the debate argued that purpose was needed as a organisational principle, or for theological reasons. Both sides of the debate agreed on a broadly mechanistic view of physics, that is to say that physical processes were best described by the idea that objects were subject to fixed laws eternal to them. As can be seen from the state of the modern world, the first side was judged to have won the debate, and all notion of purpose was exorcised from physics. In philosophy, equally, the notion of final causality was removed by Descartes, and continued in all his successors. But note the mistake that was made: the discussion of final causality only addressed the fourth of these five meanings; the first was not discussed at all except by the churchmen who

[16]Contrary to Hume, and other modern philosophers, who treated powers as though they implied certainties.

continued to follow Aristotle and were not listened to by either side. However, on account of this debate, both the fourth *and first* senses of final causality were eliminated from both scientific and philosophical discussion. There was no reason given to eliminate the first sense of final causality from the new philosophical frameworks: it was simply not discussed by the early modern philosophers; and was consequently removed from philosophical constructions without any reason. The philosophers eliminated all the senses of final causality from their world-views based on arguments directed against only the fourth sense. But because of the confusion between the different meanings of the phrase teleology was nonetheless cast aside.

This was a crucial mistake; the fallacy is that it is a confusion of categories. It is impossible to have efficient causality without final causality in an ordered universe, since if A is the efficient cause of B, then we have to state that B is a final cause of A; the two go together and you cannot have one without the other. Without final causality, we cannot explain order in the universe: why only certain effects are possible. Science itself survived this mistake; it does, after all, have experiment to fall back on, which continued to guide it in the right direction. In philosophy, however, the error was critical. It was adopted by all Descartes' successors, and has led to much of what is wrong in modern philosophy. Even today, one of the commonest arguments I see against the Thomist variation of Aristotle is that science proves that there is no need for purpose in the universe – when Thomistic metaphysics does not start by presuming that there is.

The words 'start by' above are important. Thomists do believe that there is purpose in the universe, but this is not a premise but a conclusion. It is inferred from various other premises, including the existence of final causality (in the first sense). However, this is an additional argument. Physics, for the Thomist, just requires identifying efficient, material, formal and final causes in the first sense. One does not need to bring in purpose to get an idea of how the universe evolves (although purpose is required at a deeper level to help explain the four causes and the nature of matter). Newtonian science, of course, concentrates on material causes and something with a vague similarity to efficient causality (although it can be trivially re-expressed in terms of efficient causes; the difference is that the standard discussion of Newton uses event-based causality, while Aristotle uses a substance-based causality). One could of course bring in purpose into physics, and Thomism would cope with it, but it is not necessary. Thomism does not need purpose to describe physics. This makes the debates over final causality in the early modern period somewhat irrelevant. They might have dismissed purpose from science, but they did not dismiss final causality, and if the conclusion was that science could have proceeded without any mention of purpose, then that is a conclusion with which Aristotle (or the version of Aristotle seen in Aquinas' interpretation) would have agreed. The problem is that because modern people don't carefully read the medieval writers, but only accept the opinions of their Renaissance forebears, they make the same category error that the likes of Bacon, Locke and Descartes did when they led philosophy astray.

Another important distinction to make is whether final causes are internal to a being or external to it. Does an electron in an excited energy level of an atom emit a photon rather than a bunch of flowers because it's nature is directed towards the emission of photons, or because something external to it imposes that behaviour on it (Feser, 2010)? This is one of the main differences between mechanical and neo-Platonic views of causality compared to the Aristotelian. The Platonic philosopher (and this would include thinkers such as William Paley and the deists) would both say that any directionality in physics arises from something external to the being in question: an electron emits a photon because it is directed to do so by God; the Aristotelian would say that it emits a photon because it is of the nature of electrons to emit photons. A mechanist, would, of course, dismiss any form of final causality, but their position is nonetheless closer to the Platonic position, believing that the electron emits the photon because it is bound by the laws of nature, which are again external to it. The modern discussion of teleology has focused primarily on external teleology, either

saying that an external teleology or purpose is imposed on matter or denying that; the Aristotelian position that there is a directedness that is internal to the type of being has been largely neglected.

Note that the Aristotelian intrinsic teleology does not obviously require the existence of God. Indeed, while Aristotle himself argued for God from efficient causality, he treated final causality as something basic which existed without any guiding intelligence; it is just part of the nature of the being. For an Aristotelian it is therefore possible to study physics, which is largely the identification of final causes (from which the various kinematic and dynamical properties of matter can be computed, as well as everything else studied as part of the natural world) without reference to God. There is, however, a middle way between the Aristotelian and the Platonic positions on final causes: to say that while matter has an intrinsic *telos*, God is nonetheless responsible for creating and sustaining that directionality as part of the creation and sustaining of matter. The Aristotelian states that internal final causes direct matter; the Platonist states that God acts on matter as an external final cause; the scholastic states that God directs matter using its internal final causes. Thomas Aquinas' famous fifth way is a defence of this middle position against the pure Aristotelian atheistic directionality and the Platonic external directionality. Δ

Comment:
It is worth emphasising a crucially important difference between Aristotelian causality and how the concept of causality is used by many in the modern world: Aristotelian causality is only concerned with substances, and not at all with events or processes. We can only ask what are the causes of an object. The efficient and final causes of an object point towards another object and nothing else; the material cause to underlying formless matter (which is, of course, just a idealised concept, because in practice every piece of matter adopts some form or another; but nonetheless a useful concept); the form to an abstract idea. In substance causality, it makes no sense to ask 'What is the final cause of a car accident?' or 'What is the final cause of evolution?' because neither car accidents or evolution are substances, and only substances can have final causes. Evolution, in particular, is a description of a long chain of processes which are themselves reducible to final causes.

This should be contrasted with event causality. Here the cause of an object is linked to an event; and events have causes which might be other events, or objects. For example, we may ask, "What is the cause of that broken window?" (A being.) Somebody who uses event causality would answer, "Somebody threw a brick at it" (An event). I would answer, "There was an intact window of a certain form which had among its final causes a broken window, with this potential actualised if it was struck by an object with sufficient momentum" (A substance, including a some description of its formal and final causality) and "There was a brick, of such and such a design and mass, in a state of motion with a particular momentum, which happened to be travelling towards the window." (Another substance, with a description of its form and which of its potentia is actualised.) Now, some may argue that the event based causality is a lot less cumbersome, which is certainly true. However, ultimately (if we go into sufficient detail in both cases) they convey the same information. Note that I am not sneaking in an event in the description of the final causality of the window: part of describing the final causes of a being are to list its various unactualised potentia, and under what circumstances they might be actualised. In other words, the full description would describe everything that could change the state of the window, but does not say which of these is going to occur. The event-based version of causality is specific (and makes no mention of final causes).

An important distinction between substance causality and event causality is that substance causality need not imply a deterministic universe: it is compatible with both determinism and indeterminism. Instead it is the minimum that is necessary for an ordered universe of any type. Event causality, on the other hand, has difficulties if the universe is

not deterministic. The classic example is that of a radioactive particle. The decay of an unstable nucleus is spontaneous; our theory suggests that there is no event that sets it off or hidden signal; it just happens. This causes a problem for event causality. The results of the decay (say a thorium nucleus and an alpha particle) are traced back to the event of the decay, but the event of the decay seems to have no cause. However, in substance causality there is no difficulty. A uranium nucleus decays into a thorium nucleus and an alpha particle. The efficient cause of the thorium and helium nuclei is the uranium nucleus; a final cause of the uranium nucleus is thorium plus helium, with this virtuality actualised spontaneously with a certain half life. In Aristotelian causality, we should describe the circumstances where the final causes can be actualised. This can be adapted to a deterministic system, or it can be expressed as a probability (i.e. the probability of it occurring in a certain time interval in these circumstances is ...). Thus Aristotelian causality can cope just as well with either a deterministic or indeterminate physics: the principle of causality continues to hold, so long as substances cannot pop out of nothing. Event causality, on the other hand, struggles in certain types of indeterminate physics: the radioactive decay seems to be uncaused. Aristotelian causality can therefore cope with a wider range of phenomena and possible ways in which we could have physical law: it should be preferred over other systems, to avoid putting bias towards determinate physics into our thought processes.

That is not to say that there is not an additional place for discussing the cause of events. For example, we may ask 'What was the cause of death?' 'Heart failure.' The 'Aristotelian' rephrasing of this would be something like, 'What was the cause of this recently produced corpse?' 'A living man with a heart in a state of actuality that meant it was about to fail to fulfil its purposes in the body due to fatty build up in one of the arteries causing a blockage of the blood flow; so that the living body had a final cause, likely to be imminently actualised, of a dead body.' This is obviously far more cumbersome, and too cumbersome for day-to-day usage, although (as long as we specify all four causes, and not just the final cause) both methods ultimately convey the same information. Thus I do not want to deny speaking of the causes of events has its place in common discourse; and in certain circumstance is a useful addition to Aristotelian causality. But I will not make use of event causality in this work while discussing the physics.

I should also note that causality can only (in all but a tiny number of cases) be established through theory, and an accurate and precise theory at that. We can determine causes by direct observation in only the simplest of circumstances (although, of course, observation is important in confirming theories). It is scientific theory, not experiment, which gives us the various causal links. This is, of course, why empiricists such as Hume could not understand or appreciate causality. For example, we can consider the world's climate system. If we observe the global temperatures warming, that does not explain why the temperatures are warming. It could be, of course, due to human activities such as a greenhouse effect caused by the build up of carbon dioxide in the atmosphere and enhanced by various natural feedback mechanisms, or maybe solar emissions, or less sunlight being reflected from ice caps because of an overpopulation of penguins, or maybe as part of a centuries long natural oscillation involving numerous different processes. To understand the contributions that each of the possible causes has on the fluctuations of the climate, we can only turn to a theoretical analysis. We cannot just point to a correlation between the amount of carbon dioxide in the atmosphere and temperature as though that proved anything: it doesn't. All observation tells us is that temperatures are rising: it doesn't tell us why temperatures are rising or why today's weather is linked to that of ten years ago, or how to make projections of future climate changes. Theory, on the other hand, can tell us why temperatures are rising, because it allows us to turn the various factors on and off, or change them, and find out which ones are most important and how they are related. In the case of climate, of course, the theory, when applied to the planet as a whole, is too complicated for us to do by hand, so we have to rely on computer models to make predictions from the underlying

physics. We can look at individual effects by themselves relatively simply, but once we take them out of the sandbox and into the planet itself all the individual processes interfere with each other creating numerous feedback effects which are far harder to quantify. So we need the computer model to solve the billions of equations by brute force. Of course, we then have to be sure that we have input all the relevant factors into the model. It is a difficult task to get it right, and even harder to be sure that we have got it right; all I can say is that I am glad that I am not a climate scientist with this responsibility. □

CONJECTURE 4.5.11. *Every object which exists (baring an immaterial will) must have either an efficient cause, or at least one final cause, or both.*

Proof:
I defined existence as being capable in principle of either or both of changing the state of another existent object or having its own state changed by another existent object. Now if it is capable of changing the state of another existent object then it must have a final cause describing its ability can initiate that change, or a will. This arises from the definitions of final cause and will: if something has the capacity to undergo a material change, then that change is a final cause; while the will is defined as any directedness except for final causality. (It may, of course, be able to do anything, in which case it will have every logically possible final cause.) Similarly, if it is capable of being changed, then there must be some explanation of why it is in its present state, since it cannot pop into that state for no reason, from the principle of sufficient reason, and it cannot have been in that state for an infinite amount of time. If there is a non-zero probability $\delta > 0$ – however small – that it changes its state in a given short period of time, then there is a probability $1 - \delta$ that it will remain as it is for that short period of time. Then the probability that it remains as it is for infinite time is $(1-\delta)^{\infty} = 0$. This only fails if the probability $\delta = 0$, but in that case the object is incapable of change, violating our initial assumption. If it comes into existence from non-existence, then that too is a change of state, which requires an efficient cause. Therefore every being capable of changing another has a final cause, and every being capable of changing has an efficient cause, and so every existent being has either at least one final cause, or efficient cause or both. Furthermore, we see that the only beings that lack an efficient cause (if any such beings exist in practice) are those essentially incapable of change, and the only beings which lack a final cause and a will (if any such beings exist in practice) are those which are essentially incapable of affecting others.

Note that this argument does not rely on any physical theory (while I used physics to motivate the principle of sufficient reason, the principle itself is ontologically prior to any physical theory: physics assumes the principle; the principle does not require physics). Therefore it is not only true for those beings which are the subject of physics (material beings), but everything both material and immaterial (regardless of whether or not any immaterial beings exist in practice). ■

CONJECTURE 4.5.12. *A complete description of any individual actual material object must describe each of these four causes (with the final cause in at least one of the first two senses).*

Proof:
An individual material object may firstly be described in terms of what it is, and secondly in terms of that it is. What it is is fully determined by its form and matter, as described above; this is sufficient (if we determine the form in enough detail, i.e. red ball rather than just

ball) to describe all its properties except its existence, as the objects existence is and can only be accidental to its form. We must therefore also account for its coming into existence (the efficient cause) and how it may cease to exist (the final cause). Since every property of the object other than its existence may be accounted for in its formal cause, there is no additional explanation required beyond these four causes. ■

Discussion:
Note that some of these causes may be redundant; for example its final cause may be derived from its formal cause. This does not mean that its final cause does not have any meaning; indeed all the causes may ultimately be derived from more fundamental principles, yet stating them explicitly is still useful when categorising the object. Redundant is not the same as non-existent or non-beneficial for our understanding. □

Discussion:
I will give a few examples:

A Δ^+ particle consists of two up quarks and one down quark (the same as the proton), but the spins of these particles are aligned so instead of being spin 1/2 it has spin 3/2 (naively, in the proton, two of the spin 1/2 quarks are spin up and one spin down, in the Δ^+ they are all spin up), and a higher mass than the proton and like the proton it has parity +1. It decays into a neutron or proton and a pion (together with some less probable decay channels, such as the emission of a photon to obtain a proton), and may be generated from collisions between pions and nucleons. So the material cause is the up and down quark fields, the formal cause that it has two up quarks and one down quark with aligned spin; the final causes are its various decay paths, and the efficient cause the particular interaction which led to the particle coming into being.

A teacher's material cause is a man or woman, the efficient cause is their appointment to a position in a school (or other educational establishment); their formal cause is the knowledge they possess of their subject and their ability to communicate that knowledge; among their final causes is the communication of that knowledge to the students. □

DEFINITION 4.5.13. **Instrumental cause:**
An instrumental cause is something which stands in the centre of a sequence of efficient or final causes. For example, we may summarise by saying that A is the cause of B which is the cause of C which is the cause of D which is the cause of E. There is, however, a sense in which A is the cause of E, and E an indirect final cause of A, with the intermediate states B, C and D as instrumental causes. △

CONJECTURE 4.5.14. ***Principle of causality:*** *Every change in state requires an initial state that already exists.*

This is essentially saying that things do not happen without a reason; a corollary is that you cannot create matter from nothing. A more traditional phrasing of this proposition is that every potentiality is actualised by something that is already actual.

Proof:
This follows from the principle of sufficient reason. Suppose that there is a change of state that does not arise from an existent initial state. This would in effect just pop out of nothing (note that the vacuum of a quantum field theory is not nothing). This, however, would not

have a reason for its existence: its appearance would be inexplicable, just a brute fact. It will thus violate the principle of sufficient reason, and my basic axioms.

In physics the principle of causality is guaranteed by the conservation of energy, momentum and the equivalence between mass and energy. Every particle has energy; if a new particle were to pop into existence without any initial state, then the total energy of the universe would increase; but this is inconsistent with physical law. A new particle can only come into existence if some other particle loses the energy corresponding to its mass. This other particle would be the initial state required by the principle of causality.

A further argument is that were changes to occur without cause, then we would observe that to be the case, i.e. objects appearing out of nothing, and so on. There is no obvious reason why it should be a unique event. Yet we do not observe this (I discuss objections to this argument arising from quantum field theory in section 11.6).

The argument so far only considers material objects changing in time. The principle of causality is a broader statement than this. We should also consider immaterial causes for material changes, and immaterial causes for changes in immaterial objects.

The argument above that material objects do not change their state is derived from the conservation of four-momentum, which in turn comes from the translational symmetry and locality of the laws of physics. The laws of physics, however, arise from something immaterial. Four-momentum is conserved because whatever is the origin of the laws of physics is usually unaffected by space and time (were nothing to be behind the laws of physics, they are just a 'brute fact', then there could be no violations in the conservation of four-momentum, however we should not assume it but consider the most general situation until we have reasons otherwise). If that origin is, however, influenced by particular space-time events, then the usual behaviour would not hold: there would be a preferred point in the universe, the symmetry would be broken, and something could be changed without a material cause. However, the change is not uncaused, as it requires the break down of the symmetry, which in turn is caused by the immaterial agent. We have not yet proved that such immaterial agents exist (and they may well not exist for all we have suggested so far); if they existed, then they could account for immaterial causes of changes in matter.

If an immaterial substance is changeable, then a similar argument holds. It cannot cause itself to come into being, for if it did then such events would happen frequently, which they do not. Therefore, they must be caused by something else; that something else must be able to interact with other substances and by definition it exists, therefore it is actual.

Note that the agent that causes the change does not necessarily have to be external to the substance that is changed. The change might occur due to the substance's internal final causes, or due to the actions of agents external to it. For example, a rock suspended a certain height above the ground has, in Aristotelian physics, the final cause of being a rock suspended a slightly smaller height above the ground. This change comes from its own internal nature (in Newtonian physics, we might think of a body in inertial motion in the same way: a body with a certain momentum has as its final cause a body with the same momentum but a slightly different position). The efficient cause of the rock at the smaller height is the rock at a larger height; they are the same rock, but with different potentia actualised. However, if we throw a rock in the air, it moves upwards (different potentia are actualised), because of the action of an external agent. The efficient cause of the rock at a greater height is the combination of the rock in its original place (the rock in a different potentia), and my arm as I throw it.

Note that this principle does not say that everything is caused: only that which changes in its substance or is changeable is caused. It may be that there is something which does not or cannot change its substance. If so, the principle of causality does not apply to this being: it may be uncaused in the sense that it never comes into existence because it cannot not exist.

I will return to this subject in a later chapter (11.6). ■

CONJECTURE 4.5.15. ***Principle of Proportionate Causality:*** *The principle of proportionate causality states that possible final states after a change in state are given by the final causes of the initial state; anything that is not a final cause of the initial state cannot be a possible final state.*

This states that what may be in a moment in the future is fully dependent on what is now; and by induction what may be at every future time depends on what is at the present time.

Given an initial state A, we may expect final states $B_1, B_2, B_3, \ldots$ but the final states $C_1, C_2, C_3, \ldots$ are impossible, with which states are in category B and which in category C determined only by the nature of A. We can say that A has the inherent power to become B but not the power to become C; alternatively the effects are in some sense contained within the cause.

Proof:
The proof follows from the various conservation laws inherent in nature; energy, momentum, electric charge, angular momentum and spin, and so on, as well as the possible interactions given by the Lagrangian. These explain, for example, why an up quark with a particular spin and momentum has the power to emit a W^+ Boson of a particular momentum and become a down quark of a particular momentum, while it lacks the power to decay into a pineapple – this is not just unlikely, it is impossible. Thus certain end results are possible, while others are impossible, and these are determined by the nature and properties of a particular particle. The possible end results of the interaction are determined by the initial state, and the forbidden end results are equally determined; the information needed to list the possible final states is contained within the form, matter, and accidents of the initial state.

That this list of possible final states should be identified with the final cause follows from the definition of the final cause. ■

Discussion:
I have discussed these principles in terms of physics, however they are most properly part of metaphysics. That something cannot be created out of nothing (nothing here does not imply a quantum vacuum state but the absence of any sort of actuality, material or immaterial, including the absence of space-time and any physical law) is a general principle which holds for immaterial substances (if any exist) as well as the material substances I have been discussing. The argument is similar to the material case; if something could come into being for no reason then we might expect this to happen anywhere and always, which would ultimately lead to a break down in the rationality of the cosmos (in the 'instant' of creation there would be both something and nothing, which is a contradiction). Since the cosmos appears to be rational (indeed, we have to assume that it is) the principles of causality must hold even for immaterial actualities. □

CONJECTURE 4.5.16. *If B is that final cause of A which happens to be actualised, then A is the efficient cause of B.*

Proof:
This follows from the definitions; the final cause creates a change in the state of a particle

(for example its creation); the efficient cause is that which caused the state of a particle (for example its creation). The result follows. ■

CONJECTURE 4.5.17. *The act of the will is the efficient cause of a being (or the state of a being).*

Proof:
The will is the part of the intellect that is the equivalent of the final cause in an immaterial being; it directs the intellect towards some purpose. If that purpose is related to the change of state of a material being, then the effect of the will is to create that change. Therefore the efficient cause of the change of state is the will. This change may, of course, be mediated by various instrumental causes. ■

Discussion:
Note that this causation may be indirect and instrumental; we may say in practice the will causes nervous signals to the muscles, which causes the muscles to contract, which causes the arm to move, which creates the desired effect. The will may still be thought of as the efficient cause of the change as the start of this sequence.

Equally, if an animal will is seen as no more than a certain distribution of electric impulses in the brain, as a materialist would hold, and only animals have such a will, this does not affect the argument. The will may be explained in terms of these impulses but as long as it is held to exist it is still the efficient cause of the change; one would say that some activity in the brain is the efficient cause of the change in the will to make it desire the effect, which is the efficient cause of the final effect. If the will is only partly or not at all explained by these impulses, as certain non-materialists will hold, then it will have other efficient causes. □

CONJECTURE 4.5.18. *The efficient cause of a being proceeds from either the final cause of a non-rational material being or the will of an immaterial being.*

Proof:
The efficient cause of a being (or the change of the state of a being) must have its origin in something else; which, since there are no other options, must be either material or immaterial. If it is caused by a material object, then the efficient cause will proceed from a final cause of that being, since the final cause is that which causes a change of state. If, on the other hand, its cause is immaterial, it must proceed from a will, because the will is the aspect of an immaterial being which seeks the change of another being. ■

Discussion:
I should also mention the Aristotelian concept of motion. For Aristotle, motion meant the actualisation of a potentia. What we think of as motion is movement from one place to another. However, strictly speaking, this locomotion refers to at best just one particular type of motion. I am not even convinced that we can go this far. One could (I am not convinced that one should in classical physics, but this makes more sense in quantum physics) define the inertial state of a particle as its fundamental state; thus the different potentia describe states of different momenta. In this case, movement from place to place does not involve change of state, and therefore does not involve the actualisation of a potentia. In this sense, locomotion is not even a type of motion in the traditional sense. But obviously, Aristotle

himself knew nothing of inertia. For him locomotion did involve change of state. When applied to a Cartesian metaphysics, where only extension and position are important to physics, but before the importance of the principle of inertia was fully realised, it is easy to see how the concept of motion was reduced to that of locomotion.

However, the concept of motion is wider than either locomotion or even acceleration (change in momentum). It can, for example, refer to the movement of an electron from one energy level to another, a change in the polarisation of light, and so on. All of these we would not (as moderns) think of motion, but are so in the strict sense of the word. The closest modern word to the classical concept of motion is 'change', but this does not fully describe the idea. □

CHAPTER 5

Aristotelian Physics

5.1. Overview

In many respects, Aristotle was the first physicist, in that he was the first to come up with a truly systematised and comprehensive theory of physics. Of course, he was not the first to think about the ordering of the natural world, and he borrowed much from his predecessors, not least his theory of the four terrestrial elements. We usually think of Aristotle as a dead end; I don't think anybody takes his physics seriously these days. However, it is still worth briefly reviewing his work, because it was various attempts to overcome the limitations of Aristotle's system that led to the birth of modern physics. Aristotle's physics is divided between several works, *On the heavens*, *On Generation and Corruption*, *Meteorology* and of course his *Physics*, from which the modern science gets its name. The best medieval commentary on the physics was that of Thomas Aquinas. Good modern presentations and discussions of the physics can be found in (Sachs, 1995; Leunissen, 2015; Rovelli, 2015). My own discussion here is intentionally far more limited.

The first thing that I as a modern physicist notice while reading Aristotle's works on physics is that it is not mathematical. When he discusses speed and motion this is a particularly glaring omission: he does not seem to realise how much easier it is to express his ideas algebraically. However, there is a reason for this (in addition to the undeveloped state of mathematics in his time). The centre of Aristotelian physics is his theory of causality, and causality is not something that can easily be expressed mathematically, in that it cannot be easily be measured and expressed as a number in the same way that location, extension or speed can be expressed firstly geometrically and then (through Descartes' coordinate geometry) numerically and consequently algebraically. Aristotle's physics is thus a combination of his theory of causality and various ideas about matter that he inherited from his predecessors. We immediately see that if his conclusion was wrong, then at least one of his premises must be wrong or his reasoning invalid. That Aristotle's physics is wrong does not necessarily mean that his metaphysics is wrong: logically it could mean that either his metaphysics is wrong, or his reasoning was invalid, or his ideas concerning the nature of matter were incorrect, or a combination of these. We know, of course, that his ideas concerning matter were incorrect. Thus from the failure of his physics, we cannot draw any conclusions about the success or failure of his metaphysics. This should be contrasted with modern philosophy, which is built on Newtonian physics (or, in the case of Descartes, assumptions that were incorporated into Newtonian physics). Newtonian physics is the premise, the philosophy is the conclusion. So if Newtonian physics is wrong, which it is, that is far more fatal to modern philosophy (baring, perhaps, Descartes', which preceded Newton) than the failure of Aristotle's physics was to Aristotle's philosophy.

Aristotle believed that there were five elements in nature, with the different material substances made from different combinations of these elements. Four of these (which he inherited from his predecessors, most notably Empedocles, Thales, Heraclitus and Anaximenes), air, water, fire and earth, combined to make all the earthly and atmospheric substances in various quantities. Fire is seen as hot and dry, water as cold and wet, air wet and hot, and earth cold and dry. Since there were only these two fundamental properties, any more than

these four elements would be redundant. In addition Aristotle introduced quintessence, which was the substance from which the planets and stars were said to have been made.

Each of the elements was directed by its final cause. Earth for example had the natural resting place downwards, towards the centre of the universe. Water was also directed downwards, but less strongly than earth. Air moves upwards, but less strongly, and fire upwards but most strongly. Once again, four elements are necessary to explain both the different directions of natural motion but also the differential in the strength of this movement. This view of gravity required that the earth should be spherical[1], something which Aristotle proved in part by observing the shape of the Earth's shadow on the moon, and noting that only a sphere gave a circular shadow from every orientation. The planets, however, engaged in circular motion around the centre of the universe. This is not a property of any of the four elements, so required a fifth substance, namely quintessence.

He divided locomotion in natural motion governed by the object's final causes, and violent motion, which was provided by other sources (such as a man lifting a weight against its natural tendencies), and denied that these could happen at the same time. Under gravity, bodies would come to be at rest in their natural positions.

Objects would, if there were no obstructions, move towards their natural places at a constant velocity. So if you release an iron ball, it naturally falls downwards, because it is predominately earth, and the final cause of earth is to rest at the centre of the universe. The speed of the descent depends on how much the body is made of earth, which relates to how heavy it is, so heavier objects fall at a faster speed. It also depends on the density of the medium that it is falling through: if the ball falls down, then also the medium through which it travels has to move upwards, so earth would move faster through air than through water, since water has a lesser tendency to move upwards, and therefore the total force is weaker than if the ball were falling through air. Thus the speed of the motion is inversely dependent on the density of the medium. Both of these qualities of the motion Aristotle checked through experiment.

Now, the idea of constant velocity seems odd to us: we naturally think of Newton's laws as gravity leading to a constant acceleration. And Galileo showed that the acceleration is independent of the mass. However, Newton's law is not actually what we observe in nature. In practice, we also have to take into account the friction, either the air resistance or water resistance (Rovelli, 2015). What happens is that initially the velocity is proportional to time, the behaviour we expect from a constant gravitational acceleration; but as the velocity increases, so does the force of friction. Eventually, the force of friction cancels out the force of gravity, and the body descends at a constant velocity. This velocity depends on the density of the medium, which controls the force of friction, and the mass of the object that is falling, as well as the shape of the object. To a reasonable approximation, the relation between the terminal velocity v, mass M and density ρ is $v \propto \left(\frac{M}{\rho}\right)^{\frac{1}{2}}$ (obviously there are also other factors involved, such as the surface area, type of surface, and shape of the falling body: this is only intended as a first approximation). This is, of course, exactly what Aristotle suggested. Once it achieves this terminal velocity, it will continue to descend at a constant velocity. What we see then is that this aspect of Aristotle's physics is correct in a certain limit; namely it describes the force of gravity acting in a body in a fluid after it has been falling for a certain time. For a time period during the initial descent, this relationship breaks down; afterwards it is valid as a good approximation (especially since Aristotle was fairly imprecise when it came to the details). In air, we mostly observe the period of initial accceleration unless dropping from a great height: this is why Galileo's famous experiment[2] of dropping two balls from the tower of Pisa succeeded, and confirmed the ideas he had

[1]The idea that people believed the earth to be flat until the Renaissance is a nineteenth century myth.

[2]Though it is uncertain whether Galileo performed the experiment, and if he did perform it he was certainly not the first.

adopted from Bradwardine's speculation that gravitational acceleration was independent of mass. However, the motion in air is too fast to make accurate measurements with the technology available in ancient Greece. Galileo overcame this by rolling his objects down slopes, but Aristotle did not think of that. Instead, Aristotle slowed down the descent (to a speed which could be measured) by performing his main experiments in water rather than air, or by studying objects such as leaves and feathers which have a large surface area to volume ratio, and thus reach their terminal velocity quickly. Here the friction is far stronger in relation to the gravitational force, the initial time period when his physics broke down far shorter, and Aristotle's physics describes what he saw exceptionally well.

In other words, Aristotle's theory of falling bodies is a reasonably accurate theory of bodies falling in a constant gravitational field in a dense fluid for time periods much longer than the time needed for it to reach its terminal velocity. And within these limits, it is a good approximation. And of course, this is the environment in which we live. So, we post-Newton people are often rude about Aristotle, but with less good reason than we might like. For Newton's gravity is also only an approximate theory; a much better approximation that Aristotle's, but still only an approximation nonetheless. It breaks down in strong gravitational fields (general relativity), velocities close to the speed of light (special relativity) or when energy times time is of the order of Planck's constant (quantum mechanics). It works perfectly in the limit that Newton's constant G is negligible, the speed of light is infinite, and Planck's constant is zero. Aristotle's theory of falling bodies works perfectly when the frictional force becomes particularly large.

So it is wrong to think of Newtonian physics as something entirely new from Aristotle. Rather, we see that physics has gone through a number of successive improvements and refinements, of which Newton, Galileo and company provided the first, but were themselves overthrown by relativity theory and quantum theory. Aristotle, not Newton, was the founder of physics.

5.2. Failures

Of course, Aristotle's physics had an number of important limitations, and it was these that spurred the birth of modern physics.

(1) Probably of most importance was that he failed to recognize the importance of acceleration. However, that did not lead to so obvious experimental consequences.
(2) A second drawback is that he failed to recognize the importance and significance of relative motion.
(3) His theory of circular motion of the planets around the centre of the universe (i.e. the centre of the earth) failed. This was apparent even to the naked eye observations of the classical period. The problem was resolved (at least until the invention of the telescope) by modifying the system. Instead of having the planets adopt circular orbits around the centre of the universe, they exhibited circular motion around points which were in a circular orbit around the centre of the earth; these first orbits were known as epicycles. This system, described in full detail in Ptolemy's *Almagest*, was complicated but worked exceptionally well, until Galileo disproved it through telescopic observations of the phases of the planet Venus.
(4) When we throw an object in the air, it continues to rise for a short while, before falling. According to a naive view of Aristotle's system, it should start falling as soon as it leaves the hand. Aristotle resolved this problem by saying that the motion of the arm also set the air into motion, and it was this pressure from the moving air that continued to make the body rise. This artificial upwards motion gradually expired, and when it came to an end, the body began its natural motion of descent at a constant velocity (since a body could not undergo both natural and artificial motion simultaneously). However, many of Aristotle's commentators found this explanation unsatisfactory.

The most advanced of these philosophers, and the one who came closest to discovering modern physics, was the Egyptian sixth century scholar John Philoponus, writing a thousand years after Aristotle, who resolved it by suggesting a theory of impetus, which suggests that bodies acquire their own motive power derived from the action of the thrower. This motive power gradually extinguishes over time, until it dissipates and the body reverts to its natural motion. However, Philoponus' scientific achievements did not leave a lasting legacy. He was largely ignored in Constantinople because of his heterodox theological views (he was, as all Egyptians of his period were, a Christian), and his school in Alexandria was destroyed during the Islamic invasion of Egypt. A few of the later Arabic writers paid some note to him, but he did not gain an especially large following before Islamic philosophy began its collapse. He was unknown in Western Europe until after his work had been independently reproduced and improved upon by scholars in Oxford and Paris.

However, before moving on from Aristotle's theory of projectiles, let us see what Newton says about projectile motion in a rare fluid such as air. I will use a initial approximation where the friction is proportional to the particle's velocity with the constant of proportionality dependent on the density of the air, the effective surface area of the projectile, and inversely proportional to the projectile's mass (in practice this model is a bit too simplistic, but it is close enough for my purposes). I will label time as t, the horizontal velocity at time t, $v_h(t)$, the vertical velocity $v_v(t)$, the coefficient of friction μ, the gravitational acceleration g and the mass m. The equations of motion for this system are

$$\begin{aligned} m\frac{dv_v(t)}{dt} &= -\mu v_v - mg \\ m\frac{dv_h(t)}{dt} &= -\mu v_h, \end{aligned}$$

which have the solution

$$\begin{aligned} v_v(t) &= \frac{(\mu v_v(0) + gm)e^{-\frac{\mu}{m}t} - gm}{\mu} \\ v_h(t) &= v_h(0)e^{-\frac{\mu}{m}t} \end{aligned}$$

At small times (much smaller than $\frac{m}{\mu}$), the exponential factor $e^{-\frac{\mu}{m}t}$ will be close to one, and the velocity will resemble its initial velocity, $v_h(0)$ horizontally and $v_v(0)$ vertically. At very large times, the exponential factor $e^{-\frac{\mu}{m}t}$ will be close to zero, and the projectile will seem to have no horizontal motion, and will move downward with a constant velocity $\frac{gm}{\mu}$. At very small friction, we can expand the exponential in powers of $\frac{\mu}{m}t$ and recover the usual zero friction rule that the difference between the current and initial velocity is acceleration times time.

Now let us transpose this equation to ancient Greece, and in particular a battle against the Persian archers. Two groups of people experience the arrows first hand: those who fire them, and those who are fired at. Those who fired the arrows see them moving away with the motion transferred to them with the bow. Those who are being fired at see the arrows descending vertically with a constant velocity. Both observations are in agreement with Aristotle's physics. The part in the middle of the trajectory, where Aristotle's description breaks down, is much harder to observe in detail (particularly when you are in the heat of battle). In short, Aristotle's theory was largely accepted because it agreed well with the observation of projectiles; the differences in its conclusions (in this case and other similar ones) from Newton's theory were too small to be measured. Of course, there were other situations where the differences between Aristotle's description of motion and that

> of Galileo and Newton are more pronounced, and on the basis of these we reject Aristotle. But Aristotle's physics worked much better than any rival until the time of the Medieval mathematical physicists. It seemed to agree with observation; and on that account it was widely accepted
>
> The pendulum is a more extreme version of the problem of projectiles in Aristotle's physics. In Aristotle's system, it should not rise again after falling to its lowest point. This was soon realized by scholars, and remained a fundamental unsolved problem of Greek physics, and a crucial clue that Greek physics was wrong. It was again solved by the theory of impetus.

It was the problems of projectiles and pendulums which lead to the diminishing of Aristotle's physics and the rise of modern physics. Aristotle provided solutions, but these were not seen as satisfactory. It was as various natural philosophers worked on these difficulties, and tried to find a better solution, that new ideas started to emerge, and those new ideas would lead to a revolution.

So what of Aristotle's legacy? There is a tendency in the modern era to dismiss him, and to say that every advance in the modern world was achieved by overthrowing one of his tenets. Aristotle, according to this narrative, ruled for two thousand years, and was never challenged, which led to a mental stagnation that was only overcome with the renaissance. I personally feel that this is unfair. Firstly, Aristotle did not rule for the two thousand years that elapsed between his own era and that of Copernicus. In the classical world, he was seen as just one of the philosophers in the Platonic tradition. Christianity in particular emphasised the Platonic world-view above the other schools of Greek philosophy, but the battle then was not Plato against Aristotle, but Plato and Aristotle together against the Stoics, Pythagoreans, Cynics, Epicureans, and the numerous other schools of Greek philosophy. Plato was king, Aristotle merely a courtier, and not seen as an especially significant one at that. It was not until later; possibly the late Christian period in the middle east, or possibly the early Islamic period, that Aristotle was really distinguished against Plato, and it was certainly not until the Islamic period that the Christian, Jewish and Muslim philosophers in the Caliphate during the brief period when philosophy in the Islamic world blossomed that it was generally realized that Aristotle's system was superior to Plato's. However, Islam turned against reason from the eleventh or twelfth centuries, the great philosophers were outcast as heretics, and his legacy forgotten in that part of the world.

In Western Europe, Plato, as mediated through Plotinus, Augustine and Boethius[3], was King: the only writings of Aristotle that were known until the eleventh century were his works on ethics and logic. Then, in the eleventh century Western Europe encountered both Aristotle and his Muslim and Jewish commentators, through Latin translations of Arabic and Greek manuscripts from land regained from the Caliphate and Byzantine Empire in Spain, Italy and Sicily. It took a while for Aristotle to be accepted over Plato, and he ruled for perhaps three hundred years before he started to be questioned. So Aristotle did not reign for two thousand years, only for about three hundred years in the Caliphate, and about

[3]Note that when I discuss 'Plato' in this context, I am using him as a shorthand for the Platonic tradition, that started with Plato, and developed from him, but was not static. It was this tradition that the early Christian scholars, starting from the second and third centuries, interacted with as much as the direct writings of Plato. The early Christian philosophy should be thought of as Neo-Platonic rather than Platonic, and its starting point was closer to the developed views of Plotinus than the raw original of Socrates' disciple; and obviously the Christian philosophers themselves, most importantly Augustine, developed these views and moved them away from the Plotinus in many subtle directions. However, to keep things simple, I will sloppily refer to this tradition as 'Plato' rather than digress into unimportant (for my purposes), but more precise, detail. Christian Platonism is just one of the many philosophical schools that originated with Plato, and accepted his general framework, but made subtle changes and important clarifications to many of his arguments, definitions and axioms. It should not be confused or identified with the Platonic tradition as a whole.

three hundred years in Catholic Europe, and even in these periods his rule was disputed by the anti-rationalist impulses of Islam on the one hand, and by first Plato and subsequently the early mechanists and nominalists in Europe. He was, however, the main adversary of the early Renaissance thinkers; and it was from them that the modern world did, until recently, inherit a negative opinion of the Aristotelians.

Secondly, it is wrong to say that the Aristotelian system was static. That should be obvious from the people who used him. Aristotle was a pagan; the Jews first of all took his writings, and adapted them to turn him into a third Moses (the second Moses being Plato); the Muslims adapted the Jewish Aristotle to make him Muslim; the Christians developed from the Islamic Aristotle and refined his thought further, and made him a Trinitarian. Each one took what they were given, and developed the Macedonian's thought further towards its logical culmination. The Medieval writers were respectful of the old authorities, and highly valued them, but were not afraid to develop beyond them. Aquinas was a disciple of Aristotle, but improved and developed the work of the Philosopher and the Commentator immensely, adding more arguments and making the thought more rigorous. Duns Scotus and William of Ockham developed Aristotle's ideas in different directions, and provided the bridge towards modern philosophy. In physics, as we shall see in the next chapter, there was a gradual but constant development of Aristotle's thought from his rediscovery until his abandonment. The scholastics were not afraid go beyond their master, and real progress was made, until two disasters nearly ended it: the black death, and then the renaissance (which, by turning to classical writers outside the Platonic tradition, turned its back on the medieval refinement of classical philosophy).

Aristotle survived for so long not because of an intellectual stupefaction, but because his system worked incredibly well. Not only in the experimental sciences of physics, astronomy and biology, but in ethics, and in how everything tied together. Aristotle's system was unified; today we think of ethics, physics, biology, politics, logic, metaphysics, rhetoric, poetry and so on all as separate fields fighting against each other, each with their own methods and underlying philosophies, and each largely ignoring the others. Aristotle united them; in his day there were not the fragmented sciences, but just a singular science which covered and explained everything. It proved to be incredibly difficult to beat; he needed refinement, and that was performed, but his basic concepts seemed solid. Indeed, with the failure of Newton, we can see Aristotle as just the first in a sequence: scientists who take a huge step forward, but needed a revolution in experimental precision to prove them wrong. With Aristotle, that revolution was provided by the invention of the telescope; the first glimpses of the failure of Newton came at the start of the twentieth century from various sources. We owe a great debt to Aristotle, but he was not perfect, and the world needed to move on from him. But that revolution began in an area which I don't think anyone would have predicted. It began with a small group of mathematicians in an obscure Oxford college.

CHAPTER 6

The Emergence of the Modern World-View

6.1. Definition of Science

It is my belief that there have been broadly speaking three distinct stages in the development of physics.

DEFINITION 6.1.1. **Science:**
I use the word science in its modern sense of a study of the workings of the natural world based in part on (usually but not always mathematical) reasoning and in part on experimentation. Science has various sub-fields, the most important of which are physics, chemistry and biology. Biology, which builds on both physics and chemistry, is the study of living organisms. Chemistry, which builds on physics, is the study of molecules, polymers and their manufacture and uses. It takes the basic building block of the atom, and studies things which are constructed from numerous atoms (except crystalline solids, whose study remains a branch of physics). Physics discusses everything else. Physics itself is a broad field, with many sub-disciplines, the most important of which are: fundamental physics, which investigates the most basic particles we know of (quarks, gluons, leptons, the W and Z particles, the photon and the Higgs Boson), partly studying how these combine to make the next level of particles (which includes the protons and neutrons that make up an atomic nucleus) and partly trying to reach the level that underlies these particles; condensed matter physics, which looks at the theory of solids and crystals and their properties; the physics of fluids: gases and liquids; nuclear physics, which investigates the nuclei of atoms; atomic physics; statistical physics and thermodynamics, which studies temperature, pressure and other phenomena; optics and electromagnetism; and astronomy and cosmology, which study gravity, the early universe, and the nature of stars and planets. These are all linked by the use of similar methods and mathematical tools, and ultimately each sub-field of physics (in their modern formulations) started with insights from Newtonian mechanics, although they have moved on from these beginnings to be dominated either by relativity or by quantum physics.

One of the distinguishing features of science is that it attempts to explain the natural world on the basis of various foundational principles with, once they are accepted, are self-sustaining and impersonal. This should be contrasted with previous attempts to understand the natural world which relied on each element of it being associated with its own spirit and god. Instead, it seeks to find a unified framework which explains everything without reference to supernatural intervention. It assumes that the universe is rational and in some senses regular and predictable.

This methodology does not, of course, exclude supernatural intervention on a continual basis. Science only formulates the laws. It does not and cannot explain why they are obeyed. Rather, the laws of physics are stated without explicit reference to any gods, fairies or demons. If there are such things, then either their actions would be included within the laws of physics, or we would have to say that the 'natural' regularity of the universe is sometimes broken by such things. What do I mean by saying that their actions are included as part of the laws of physics? As an example, if their actions in moving observable matter

were regular, then those actions would be incorporated into scientific law without any explicit reference to the being performing the actions. The scientist would be describing the works of the fairies without ever realising that that was what he or she was doing. In either case, physics by itself cannot rule such creatures out or rule them in: it simply, by its choice of methodology, has nothing to say on the subject. The purpose of physics is to describe the fundamental principles that underlie the material world. Description is not the same as explanation. Physics, and science in general, does not seek to explain those principles that is the job of the philosopher or theologian. The philosopher can, of course, legitimately postulate fairies, goblins or gods if they seem to be a viable explanation of why physical law takes the form that it does. Admittedly few philosophers would go down this path. Any who do would need to give a detailed explanation of why their particular goblin would imply the effects that we observe. Equally, anyone who proposes a purely naturalistic explanation for the supremacy of scientific law needs to back up that proposition with evidence.

This should be contrasted against the more ancient use of the word, where 'science' indicated all knowledge, so it also included, among others, philosophy, mathematics, theology, ethics, politics, history and artistic expression. △

DEFINITION 6.1.2. **Greek science**:
Greek Science was the first organised, systematic and reasonably successful attempt to study the natural world without recourse to random interventions from various spirits or gods. It was based on experimental observation and logical (but not, on the whole, mathematical) analysis. Classical science emerged from considerations concerning the drawbacks of Greek science, as quantum physics arose from discussions about the limitations of classical science. △

DEFINITION 6.1.3. **Chinese and Indian science**:
The Greeks were not the only ancient society to develop a science. Other sciences, in some respects similar to the Greek, were established in both China and India. However, these systems failed to develop and had little lasting effect on world history (although the underlying principles of ancient Chinese science are still important in China, where, for example, they support Traditional Chinese Medicine. However, these principles are contrary to those of modern science, and have thus been shown to be false by the countless experiments that support modern science). However, the early Indian mathematicians did leave a lasting legacy, and many important technological breakthroughs, inspired by their sciences, were achieved by both the ancient Indian and Chinese societies. △

DEFINITION 6.1.4. **Classical science**:
Classical science is the system based on the work of Newton and Galileo. It is distinguished from Greek physics in that it treats the natural world using mathematics rather than the other branches of logic, and because its experimental background is conducted more rigorously. It is distinguished from quantum physics because it is based around the deterministic mechanistic system rather than the probabilistic and indeterminate contemporary constructions. I also include special and general relativity in classical science; although special relativity is fully absorbed into quantum physics, and the theory of quantum gravity (which we don't yet have) will merge general relativity with quantum physics.

There is, of course, a slight issue of terminology. The word 'classical' is usually used to denote Greek and Roman culture, and the early medieval philosophies which were based on it. We should really use the word to describe Greek science rather than Newton's, and find some other way to describe the subsequent developments in physics. However, unfortunately this terminology is now too common and we are stuck with it. Δ

DEFINITION 6.1.5. **Quantum physics**:
Quantum physics is science based on the principles of quantum mechanics and (later) quantum field theory. Although it incorporates many of the ideas of classical science, its foundations are very different. Classical science emerges from quantum physics in the limit that Planck's constant $\hbar \to 0$ (so if we were to expand the results of quantum physics in powers of $\hbar$, the leading order term in the expansion would give something equivalent to Newton's result, with the remaining terms corrections to it). Δ

I will use the term 'modern science' to describe both classical science and quantum physics.

In this work, I will concentrate on physics, the most fundamental of the sciences, and the most fundamental part of physics at that. My reason is that I am interested in the links between science and philosophy. That linkage exists first and foremost between metaphysics and the most fundamental parts of physics. I will therefore concentrate my attention there.

Please note that although my discussion might seem to be historical, it is not. I am a scientist, not a historian of science, and like most scientists I make a bad historian of the subject, simply because my interest is different. My primary purpose when looking at an old source is 'How much of this is true? How did this contribute to the standard presentation of physics today? Is there anything in here whose importance is only realized when looked through the lens of modern physics, so I can turn this overlooked nugget into a modern research paper?' A competent historian, however, is not interested in the truth of the ideas in itself, and what future generations made of those ideas is only discussed as a motivation, if at all. Instead, they discuss how these ideas were presented and debated at the time, turning to the writings of even those we would today would regard as minor players but who gave an important contemporary commentary, and how the ideas were built on the older authorities and conceptions. My discussion, based on (Hannam, 2010; Grant, 2006, 2010; Lindberg, 2008; Burtt, 1931), is thus limited as a history, though, I hope, useful as a discussion of the principles that led to the emergence of the modern world.

My intention is not to give a history of ideas, but to explain the modern way of stating those ideas, and to examine the basic assumptions upon which those ideas were established. I need to study the assumptions because my intention is to challenge this narrative, and the best way to challenge something is to show that its assumptions need not necessarily be correct. The study of assumptions requires going back into history to the people who first made them, and seeing when each idea was proposed and in response to which challenges. However, because I am interested in the modern description of those ideas, as it was standardised in the nineteenth and twentieth centuries[1], my presentation will be full of anachronisms. In some cases I treat the ideas out of historical order. On first reading, Newton probably would not recognise my description of his work, and certainly not on a superficial reading. Only when going through the details, and translating from my language

[1]It is often the case that an initial genius is stuck between two different world-views: that which he is responding to, and that which he is establishing. It then falls to later scholars to fully systematise his ideas in a way fully consistent with his principles. Sometimes this process takes centuries. It is then the systematised presentation, not the original presentation, which is presented to students and taught. It is the genius' thoughts, but presented in a more logical and self-consistent way, while the original formulation of those thoughts can often be quite disordered and messy as one by one the various pieces fall into place.

to his, will he agree that I am describing what he did. Whether he would accept that the modern way of presenting classical mechanics is superior to his own is something we will never know. Some mention of history is necessary for my purposes, but I am not writing a history here.

6.2. The Laws of Physics

One problem we must face first is what do we mean by laws of physics. I have used the expression myself already, and it is in common parlance. The problem is that the same phrase is used to mean different things in different philosophies, so it is worth outlining some of these uses before I begin. I will give my own understanding later. My list is not intended to be exhaustive (for example I skip over Plato's definition), but only to highlight various understandings which I have or will encounter. Nor do I claim that these categories are precisely determined: there are numerous variations between different thinkers within each of these schools. This should just be treated as a broad summary of several distinct positions. While there are strong similarities between some of these, others are very different from each other. It is important when reading a writer to understand what he understood by physics, and not to force our own definition onto that writer. That would lead to misunderstandings. It is important to also realise, when starting to think about these things, that there are other possibilities than just the one that we hold; and that the perspective we were raised to believe might not the only one that is consistent with modern science. Most people today are, I think, either mechanists or instrumentalists, with a few lingering empiricists and Aristotelians.

(1) The Aristotelian vision. This focuses on the processes of efficient and final causality. Rather than a universal law of physics which everything is bound to, it focuses on individual laws for each individual object. There is a connection between causes and effects (which refer to substances alone and not events), but this connection could be accidental; it does not follow that given object A, circumstances B must follow, but only that they could follow. At best, one might say that they would usually follow if not otherwise impeded, or one might leave the consequences indeterminate, and say there are several things that might happen.

(2) The mechanical vision. The mechanical vision of a law sees it as a universal feature of the universe, to which all matter is bound indiscriminately. The law can be expressed mathematically and until recent times most of of those who held this view would have made determinism another key characteristic. The picture used is that of a clockwork machine, operating with fixed regularity, and running without outside intervention. There is thus a necessary connection between causes and effects, which may be either substances or events.

(3) The deistic vision. The vision of the laws of physics in deism is similar in many ways to the mechanical approach, but with the exception that it is understood that the concept of a law requires a lawgiver. The laws operate as part of God's design of the universe. They are thus dependent on God, but don't need His ongoing intervention, although it is possible for God to break into the universe and intervene. Rather, as God creates the universe, he also provided a law that governs the motion of matter, just as a government creates a law that (allegedly) governs the motions of its society.

(4) The empiricist vision. The empiricists did not believe that it is possible for us to perform a theoretical analysis. We thus cannot say that there is a necessary connection between two objects. All we can observe is that a particular effect regularly follows a particular cause. Thus the laws of physics are no more than a description of observed regularities. For example, one might plot the pressure of a gas at fixed volume against its temperature, draw a straight line through the data, and that straight line would be the law of physics. Someone of the mechanical or

Aristotelian philosophy would try to explain why the data was in a straight line, either in terms of mechanical forces or final causes. However, to the empiricist, such attempts are futile: either we cannot know the reasons behind the law, or there are no fundamental reasons, just the observed regularity.

(5) The instrumentalist vision. An instrumentalist sees the laws of physics as merely a human construction that don't tell us anything fundamental about reality. They are merely a tool used by physicists to make predictions. For example, if combined with idealism, the laws of physics might make connections in the mental world of human ideas, but there is a chasm between that world and that of physical reality. This view is opposed to scientific realism, held by mechanists and Aristotelians, in which physical laws are real features of nature. The laws uncovered by physicists only connect observations, rather than actual physical objects.

The main difference between instrumentalists and empiricists is that instrumentalists believe that theoretical expressions of laws (such as Newton's law of motion or law of gravity) have some use in a mental construction, while empiricists would deny that they have any use at all, and all we can rely on are experimentally derived laws such as those formulated by Kepler. The other three positions are scientific realists, which believe that there is some law of physics which is built into the fabric of the universe. The difference between the deist and mechanistic approaches on one hand and the Aristotelian approach on the other is that the Aristotelian approach views the laws as being local, that is they arise from the nature of individual substances, while the deist and mechanical approaches view them as universal, that is external to individual substances but to which those substances are in some way bound. The difference between the mechanist and the deist (as I have expressed them here) is that the deist believes that the substances were directed to obey the laws by God at their creation, while the mechanist removes God from this picture, and (for example) states that this is just how things work. I should also clarify that the mechanist conception of law was originally proposed as part of a wider philosophy, what I call the mechanical philosophy. But it was just one of the premises of the mechanical philosophy, and can survive without the others. Indeed, it has survived, because many people still hold to it even though they no longer accept mechanical physics.

There are, of course, different perspectives we can take beyond these models. Indeed, my own view, which I will discuss later, doesn't really fit perfectly into any of these categories, although it is closest to the Aristotelian. I have discussed the Aristotelian approach in the previous chapters. In this chapter and the following ones I will discuss the mechanical, deistic and empiricist views. I will conclude this section by saying a few words about the instrumentalist view.

I am of the view that the instrumentalist vision of physics is not of much use, however popular it might be among philosophers. It was introduced on the basis of two basic premises. Firstly, German Idealism, namely that there is a clear distinction between the world of our knowledge and the world of reality. Secondly, the understanding that data is not fully determining, i.e. numerous theories could be consistent with the same data. We thus have no way of knowing which theory is correct, in the instrumentalist view. Thus any conception we have can only be a human construction, and not part of reality.

Now, there is of course, some truth in this. There is a difference between the equations written down by physicists today and the true laws of nature that they are meant to represent. And scientific realists will acknowledge that. However, instrumentalists take a jump from "We cannot know the true law of physics" to "There is no true law of physics in reality." A scientific realist might not even accept the first premise, that "We cannot know the true law of physics." They would certainly accept that "We do not know it at present," and maybe "We cannot know it with perfect precision," but the word *cannot* without qualification disguises the fact that we do, already, know a great deal, if not everything perfectly precisely. The difference between scientific realists and instrumentalists is that scientific

realists acknowledge that there is a true law of physics which can be represented mathematically, and our best equations today are an increasingly accurate and increasingly precise approximation to those true laws. The laws are out there for us to discover. And it is a voyage of discovery, not a work of invention.

Precision is important in this discussion. For example, there are numerous parameters in physics which need to be measured. The best that physicists can do is say that they lie between a certain range. If the true parameter does lie in that range, then the theory is an accurate but imprecise construction of reality. Instrumentalists tend to focus on accuracy; precision is important as well. Recognising that one can be accurate but not precise avoids the main argument of the instrumentalists. If physics depended on a parameter, and to know physics we would have to know that parameter to an infinite number of decimal places, then it is true that we can never hope to achieve such precision. However, it might be enough to merely say that the parameter lies within a certain range. That can be achieved, and it still counts as accurate knowledge of the laws of physics. But for it to be accurate or inaccurate implies that there is a law of physics in reality against which our equations can be compared.

So why might we reject instrumentalism? Firstly, it doesn't explain the ability of physicists to make predictions, such as the discovery of the Higgs particle. If the laws of nature are just in the mental world, then why should the physical world be bound by them? Why should the Higgs be there in reality, just because it is needed to make our equations consistent, if our equations are not descriptive in some way of reality?

Secondly, it doesn't explain why physics is so successful at modelling nature in terms of what has been observed before. If it were merely a human construction with no connection to fundamental nature, then it would be somewhat miraculous that its results work so well.

Thirdly, scientific realism provides by far the simplest explanation of why we observe what we do. If it looks like a duck, quacks like a duck, has webbed feet like a duck, lives on a duck-pond, and answers to the name of Donald, then it is almost certainly not merely a mental image of a racing pigeon. Similarly, if our observations are precisely what we would expect if a particle of charge -1 and rest mass 0.511 MeV emerged from a cathode ray tube and hit a screen, then could it not be that there really was such a particle which made that journey, rather than we just had an unconnected series of mental images? The second set of principles of modern physics are the symmetry laws, which are used (for example) to predict the types and natures of composite objects, and to constrain the possible interactions between different particles. Symmetries can be directly observed in nature, just like particles, so while they are a little more abstract, they are no less the product of raw observation. The various interactions and forces affecting the electron on its journey are ultimately restricted by those symmetries. Modern theoretical physics is constructed from far fewer premises than those who first proposed instrumentalism in the early twentieth century could have imagined. The examples originally used to justify the philosophy are out of date. This doesn't make instrumentalism wrong, but it does mean that it would need re-justification in terms of contemporary physics. This makes sense if any version of scientific realism were correct, where we can reduce our understanding to a small number of fundamental principles taken from observation, but not so much in an instrumentalist view which denies that there ultimately can be such principles. Even if there are bits missing of our current theory – perhaps fundamental particles with masses larger than we have currently been able to reach, this doesn't deny that there is a true theory which we can, in principle, understand.

Finally, the instrumentalist view has failed in the past. For example, it was originally supposed that atoms were merely a mathematical convenience, until they were directly observed. Now the debate is over quarks and leptons. Why should we expect instrumentalism to fare any better now than when it did a hundred years ago? We should be suspicious of any philosophy that has a track record of getting things wrong. Thus the instrumentalist

understanding of physics does not explain why physics has developed as it has done and why it is so successful, while the realist understanding does.

6.3. The Emergence of Classical Physics

Classical science comes from the merger of various different ideas. Some of these ideas date back to ancient Greece and Rome, albeit with important developments from the Islamic world. However, they each individually received fresh impetus in Western Europe in the late medieval period, and by the time the black death led to the downfall of the medieval intellectual culture, they had each developed considerably from their classical ancestors. They were merged together during the early modern period, and subsequently developed at a rapid pace.

One question that needs to be asked is why did classical physics emerge in Western – Roman Catholic dominated – Europe? Other civilisations were far more advanced when the initial seeds of Western Europe were sown as it emerged from the dark ages. The principal reason, I think, is that the religious culture was amenable to it. Obviously we like to think of the Galileo affair as the beginning of conflict between church and science – and I will briefly mention this affair later – but I want to go back to the period before Galileo. Christian theology grew from three sources: the third was Greek Philosophy, and in particular the tradition founded by Socrates and Plato (which included Aristotle, Ptolemy and Plotinus); the second was the ancient Hebrew religion; and the first was the life, teaching, death, and resurrection of Jesus and his disciples[2] which extended the Jewish religion. The Hebrew religion was distinct from most around it. Firstly, it was strictly monotheistic. There were no competing gods warring over the natural world, but only one God in complete control of everything. Secondly, especially in its later developments, it believed that God is rational, and ruled the world according to rational principles; and also that God is entirely different from us, not an old man sitting in the sky driven by near-human motivations, but something wholly unimaginable. God cannot lie, and does not play tricks on us. What this means is that we can examine the world, and see something of the mind of God behind it all. We expect the patterns of the world to be regular, because there is a single unified force guiding it. Thirdly, it emphasises our own free will, and ability to control and modify nature: it is not fatalistic. Things do not just happen because of God's decrees and we are just puppets in a play already written; but God bases his decrees in part on our free choices. This means that we should try to ensure that our choices are reasonable, which requires developing an understanding of the universe. We are not fated to ignorance, but have the God-given ability to overcome our ignorance, if we choose to use it. With our minds made in the likeness of God, we have the capacity to understand God's creation. Fourthly, it emphasised that the world was regular, not a cyclic history, but history was a straight line from a beginning to an end, and ruled by understandable and fixed laws. All of these are enshrined in the Jewish and Christian scriptures. Equally, all of these ideas were derived independently in the Platonic philosophical tradition – the similarity between Plato and Christian religious thought was one of the reasons why Plato was so quickly adopted by the early church.

We can compare this to other traditions. The Indians, ancient Mesopotamians and Egyptians, and the Greeks and Romans before Constantine were polytheists – baring a few semi-heretical exceptions such as Socrates and the school he founded, they did not believe in a single unifying principle and thus had no motivation to find one. The Chinese were polytheistic in religion and dualist and fatalistic in philosophy. The Persians were dualist in both philosophy and religion. The idea that there is a single and rational unifying principle

[2] I am not yet in a position to discuss whether the resurrection happened. However Christians, from the original first century eye-witnesses to the present day, obviously strong believe that the genuine and bodily resurrection occurred and placed it as the keystone of their belief system; and on the basis of this belief – whether it is true or not – built their theology. As I am describing the main premises of Christian theology, it its legitimate for me to include it in this list.

behind the workings of the world is thus alien to these cultures. That doesn't mean that some great thinker might not propose this (as Socrates, Plato and Aristotle did); but it does mean that the wider culture would suppress attempts to take their ideas to their logical conclusion. Most other cultures either adopted the ideas of one of these great civilisations, or accepted some form of animism or polytheism without developing any real philosophical tradition worthy of the name. The Arabs (after Islam) were strict monotheists, but believed that God was outside of human understanding and knowledge, and ultimately unpredictable, and were also fatalists, denying or at least severely reducing the understanding of free will. When Islam interacted with the Christian, Greek and Persian world, it also at first adopted Greek philosophy, and came closest to overcoming the crucial hurdle that started the process of modern science, but it too failed. There is a much larger tension between the philosophers and the religious tradition of Islam than exists in Christianity or even the Persian religions. Thus the Islamic world was split into two camps; the first used philosophy and reason, and made some progress towards modern science, and the ideas developed then provided one of the sparks that helped ignite the spirit of inquiry in Europe. The second, based more on the Islamic religious 'revelations', believed that attempts to understand the universe were seen as futile, and the only sure source of knowledge was the Koran and, to a lesser extent, Hadith. There was a debate between these two camps over the course of several centuries, and the more religious side won. Rational inquiry shut down in Islamic society just as it was rising in Western Europe, and, aside from Koranic studies and Islamic law, has not been revived to this day. Whatever progress there is in the modern Islamic world (and there isn't much) tends to be in the less devoutly religious societies and institutions; the reverse of what happened in Medieval Europe. The Byzantines perhaps had the best opportunity, but also failed. Partly this was because of the Arab invasion of Syria and North Africa, which took away key lands at just the wrong moment, and leaving the Eastern Empire with just a shell of its resources, and the remnant forced to continually fight for its existence against incursions from the Bulgars in the North and repeated attempts at invasion from the Muslims in the South and East; the Crusaders dealt the deadliest blow leaving defeat inevitable, and the Ottomans duly finished off the job. An insecure society does not provide the freedom to develop radical ideas. Perhaps the rigid political structure of the Byzantine Empire was also a hindrance; partly, perhaps, there was just ill luck that aside from Philoponus no scholar of sufficient imagination was born to Byzantine parents.

One further important advantage that Western Europe had was the separation of church and state (aside from anomalies such as the papal states), and how it splintered into different political authorities. The church provided an overall unity to Western European society; it provided a common language and framework, it prized (at least at the time we are referring to) both secular and religious knowledge and encouraged study and learning, and was a stable midwife for the emerging academia. On the other hand, the fragmentation of the state meant that innovative scholars who offended one set of politicians can always move to find another university and leaders who would support them. There was a great deal of patronage of the universities by the landowners – having a great scholar in your employ was one way of showing off to your rivals. If you offended the church, you could appeal to the protection of a king. If you offended one king, you could move to his enemy. This gave the early medieval scientists and philosophers a great deal of intellectual freedom. Not total intellectual freedom: the church still raised some no-go areas. But generally, the scholars were as free as they ever have been since. Moreover, the common scholarly language, Latin, meant that there were no barriers on their movement from one place to another, or communicating ideas across borders.

The advantage of the culture being favourable to scientific inquiry is that it was necessary to produce a chain of scholars capable of making progress. Some like to point out geniuses as the hallmarks of scientific progress, with science moving from one genius to

another, with great individuals providing shifts in paradigms. A genius is capable of transcending their culture. Philosophical and scientific geniuses emerged in China, India, Persia, the Islamic world and others; they almost made the key breakthrough. But their ideas did not catch on, and died out. Science is really driven forward by a succession of competent, but no more than that, scholars. One genius can do more than the merely competent, but the merely competent come in far greater numbers. They are necessary to prepare the ground for the genius, and to tidy up the loose threads that the genius left undone, and perhaps to explore various blind alleys. The genius is an essential part of the process, but he needs to be surrounded by a crowd of the competent for his ideas to catch on and be developed. Most of these scientists are anonymous, forgotten by all but the most specialised historians, but we should never underestimate their contribution. Now the merely competent cannot transcend their culture; they plod along in the direction that the wider society they dwell in is pointing. Thus only if the culture is amenable to scientific inquiry can we find this chain of competent scholars that are necessary to build upon and refine the work of the geniuses. In early Roman Catholic Europe it was so amenable; other cultures as a rule were not, or only partially, or only for a brief time.

The dominant medieval view of physical law, articulated best by Thomas Aquinas and Albert Magnus, but not only by them, was in part borrowed from Aristotle, and in part based on the earlier Plato-based Christian theology. Aristotle characterised instead of a law of nature, laws of natures. Each individual substance had its own tendencies and powers, based primarily on its final causes. By understanding each individual final cause, one could understand how the natural world evolves. The second aspect of physical law was taken from Christian theology. God upholds and sustains the universe. He is not just the absent landlord of deism, who sets things in motion and then has nothing more to do, but is actively involved at each moment. Thus God not only gives each type of being their powers, but also guides them into fulfilling those powers. In particular, existence is seen as an accidental property of matter. There is no reason for any matter to continue to exist unless its form is continuously conjoined to an act of existence. That conjoiner is God, who (unlike material substances) is both pure actuality and pure existence. The natural powers of an object are just seen as an intermediary step, standing between the observed effect we see, change, and God as its ultimate author. In other words, if God stopped working at any one moment, then everything would immediately cease to exist. However, God does not stop working; because God is timeless, external to the universe, unchangeable, constant in His nature, and regular in His behaviour, as well as being of a single simple, unified, nature.

What this means is that God's actions to sustain the universe are (baring the odd miracle) consistent and predictable. One can understand the regularities of nature without direct reference to God; we can separate the questions of physics, 'What are the powers of each substance?' and philosophy, 'Why does it have those powers?' The second of those questions cannot be answered without mentioning God; but the first only depends on the natural substance, and can be answered without reference to God. Greek physics is concerned with understanding and enumerating the various powers of each type of substance. God is there in the background, the cause of the regularities of nature, but those regularities can be understood by themselves. Of course, they inherit all the properties of God as their ultimate source. The natural regularities do not change in time or from one place to another; they are rational and predictable and can be studied through reason; they are the same for each type of being, only depending on the nature of the substance. They are internal to the substance in the sense that the final causes are part of the nature of the substance, but external and objective on account that they originate ultimately from God. The regularities do not depend on who is observing them, and they are repeatable: if you perform the same measurement in the same circumstances you should get the same answer. All of this is guaranteed by the nature of God and the way that the laws of natures are ultimately

dependent on God. It was not obvious outside a Christian culture that these assumptions behind modern science, which today we take for granted, were true.

Of course, this also means that studying nature is a means of studying God. God revealed himself in two books; the first the Holy Scriptures, and the second the natural world. Each one of them, used properly, can give us a greater knowledge of, and closer relationship to, the Lord of the universe and Friend of man. Such, anyway, was the world-view of one important school of Medieval academia.

I shall call this view the causal view of physics, and it is best distinguished against the mechanical view by saying that the causal view believes that the evolution and changes in matter are best described by things internal to the bodies (namely formal causes, final causes and other dispositions) while the mechanical view believes that locomotion is best described by something external to the bodies (in this case fixed laws of nature) acting upon them. In the causal view, objects are active generators of motion; in the mechanistic view they are passive recipients of motion.

You will notice, however, that this view of the universe is teleological: it is built on final causes. You will also note that physics does not depend on purposes, though an understanding of God's purposes is necessary to answer the questions of philosophy.

There were also opponents to this Aristotelian viewpoint in the medieval world. In their own time, Albert and Aquinas were controversial figures, both from the Platonists, the hardened sceptics, and even the more hard-core Aristotelians (Aquinas adapted Aristotle to provide a synthesis with Augustine's thought; others were not so willing to do so), and not everyone agreed with him or the viewpoint I have just described. So, in the fourteenth century when our story starts, we had the thoughts of the likes of Duns Scotus and William of Ockham, who diverged from a 'pure' Aristotle in different directions from those chosen by Aquinas, and who provided the bridge between Aristotle and modern philosophy, but also numerous people who accepted Plato or the Neo-Platonic theories, or at least were still influenced by them. This proves significant, because while Aristotle minimised the importance of mathematics (focusing on dialectic or logic instead), Plato, particularly in his later years, became convinced that Pythagoras had been correct to say that the fundamental principle behind the universe is geometry, and he made this a central point of the curriculum of his philosophical school. After Plato's death, his academy passed to some of his former students, and they, if anything, strengthened this concept. When the work of the academy and Aristotle's lyceum were merged together by Plotinus in the Roman period to form the Neo-Platonic tradition, the Pythagorean idea that geometry was important to physics remained. Plotinus' philosophy was adopted by the major Latin Christian thinkers, such as Augustine and Boethius (whose mathematical works and translations of Aristotle's treatises on logic were key texts used in the early universities), and through them it was transmitted to the Western Church, preserved in the monasteries through the dark ages, re-emerged in the early church schools, which combined together to form the universities. The Pythagorean and Platonic emphasis on geometry was only a small part of the limited collection of Greek learning preserved through the translations by Boethius and others as the Western Roman Empire collapsed, but enough of it survived to inspire later thinkers.

On the other hand, Aristotle himself was scornful of Pythagoras' ideas, and considered Plato's acceptance of them as one of his mentor's biggest mistakes. This contributed to the division between Aristotle and the rest of Plato's students, and that division ultimately led Aristotle to set up his rival philosophical school in Athens. But it was Aristotle's school, not Plato's, which gave the ancient world the first description of physics which was anywhere close to being complete and conforming to the observed world. Thus Aristotle's prejudice against thinking about physics geometrically became the mainstream position in classical and later Byzantine, Persian and Islamic thought. However, medieval Europe did not know of Aristotle's concerns until the controversial re-introduction of his works in the twelfth century, and by then an intellectual tradition based on Neo-Platonic thought had

already been established. The debates of 4th century BC Athens between Plato's disciples and Aristotle's disciples over the role of mathematics in physics were re-enacted in medieval Europe; only this time mathematics had advanced enough to give the Pythagorean viewpoint a chance. From the eleventh century, scholars attempted to find Pythagoras' proposed link between geometry and physics, but without any progress: they lacked the most advanced description of physics, which was Aristotle's, and the most advanced mathematics, which at that time was being developed in the Islamic world based on Greek and Indian sources. But the idea that physics *ought* to be represented geometrically was hanging around in the background of the medieval universities even after Aristotle started to dominate academic thought as our story begins in the early fourteenth century. The roots of modern science thus emerged by taking some aspects of the Aristotelian world-view, and some aspects of the Pythagorean world view, mediated through Plato and Boethius, and merging them together into a new entity.

The medieval European universities arose out of the church schools that dated back to Charlemagne in the early ninth century, established because of the need to educate priests and government bureaucrats, lawyers and officials. Charlemagne's father and grandfather had beaten back the first of the great attempted Islamic conquests of Europe, but culturally and intellectually the society he had inherited and unified into a new Western Roman Empire was way behind its Islamic adversary and Byzantine rival. If the Empire was to survive against its far more powerful enemy, that had to change, and rapidly. An attempted alliance with Constantinople was rebuffed (who was this Germanic barbarian to claim the title of Roman Emperor of the West?); so Charlemagne was forced to use the few resources that remained in Latin Europe to quickly develop his Empire without outside help.

The new schools were mostly linked with the church (the only institution that prioritised and preserved learning in the dark ages leading up to the time of Charlemagne), although sometimes founded and governed by secular authorities. Their primary purpose was to train priests, medics and lawyers, and these led to the schools of divinity, law and medicine which provided the focus of the early universities. However, before graduating to these advanced studies, the students had to undergo an initial set of studies in basic and foundational knowledge. Because it believed that all knowledge would ultimately lead back to God, the church supported rather than opposed secular knowledge. Thus the students were taught to master Latin, grammar, rhetoric, logic, philosophy, ethics and metaphysics; the natural sciences of physics, astronomy and biology; the mathematical sciences of geometry and algebra; and also music and aesthetics before moving on to their specialised studies. Thus each student and scholar at the universities had a thorough and up-to-date grounding in both physics and mathematics; but they were taught as separate subjects.

Now physics was not in general seen as mathematical; there was some geometry in astronomy, of course, since the orbits of the planets were governed by geometrical figures. But, following Aristotle, the physics of the heavens was seen as wholly separate from the physics of the earth. Most physical properties were not seen as being able to be expressed mathematically, although there was some understanding of the relationship between extension and geometry. However, Aristotelian physics and Platonic mathematics (with the syllabus now improved by the Islamic mathematical texts) were taught side by side, and it was perhaps inevitable that somebody would succeed in combining them. That occurred in fourteenth century Oxford, and that moment is one of the candidates for the birth of classical physics.

6.4. The Merton Calculators and Their Successors

Precisely when classical physics began and who can be said to be its instigator is not entirely clear. There were various streams of thought in the thirteenth and fourteenth centuries, each of which contributed to the start of the process. However, the primary difference between classical and Greek physics is the mathematical nature of classical physics, so a good candidate is the first successful geometrical description of physics.

Intellectually, it was a turbulent time in Western Europe. The early universities were very much church institutions; their main purpose was to train clergy, lawyers, and doctors and give the nobility a good understanding of rational thought, which obviously had Christian theology as its foundation. Of course, with the firm belief that all true knowledge originated from God, and that secular knowledge also came to Christianity though by a different source – and perhaps give some insights not included in revelation, their syllabus was not only based on Christian sources, but the main part of it on the best of what was available elsewhere, mainly Greek and Roman sources, together with some Indian learning (particularly in mathematics) mostly transmitted indirectly through the systematisation of the Persians and then the Islamic world; though some no doubt also transmitted directly from the various Byzantine controlled areas in Italy. Platonic thought had been strongly defended (following the practice of the early church) when the church schools were first set up, and continued as they merged into the first universities.

Then, from the twelfth centuries, translations into Latin of Aristotle and his Islamic and Jewish commentators began to widely circulate, and a second stream of thought emerged, culminating in the work of Thomas Aquinas in the thirteenth century. However, there still remained a number of scholars who were convinced Platonists, and others who used the intellectual turmoil to question both of them. In particular, there was concern about Aristotle's physics, which seemed to be based neither on purely rational grounds, nor on the authority of Scripture and the early church. Nonetheless, the *Physics* and his *On The Heavens*, together with Ptolemy's *Algamest* which was built on them, were the best works of their kind available; however, this lack of surety made some question whether certain of their contents (such as that the earth was immovable, a vacuum was impossible, that the natural movements of heavenly bodies could not be in straight lines, and the universe was eternal) restricted God's freedom to order the universe as he liked. In particular, various condemnations were issued by the Church at Paris in the thirteenth century trying to restrict the teaching of certain ideas as facts. In particular the edict of 1277 issued by the Bishop of Paris condemned various aspects of Aristotle's works, no doubt trying to restrain the over-exuberance of certain scholars at the University of Paris.

This edict only had local and limited effect. It did not suggest any new method or ideas to replace Aristotle, and thus itself cannot be thought of as the positive spark that started modern science. It did, however, lead to a new sense of intellectual freedom, and encourage some scholars to think about whether there were other ways to do physics other than the method of Aristotle, to place the subject on firmer ground. Thus in the late thirteenth and early fourteenth century, while many scholars remained devoted to Aristotle's physics or Plato's conception of the universe, there were others who searched around for an alternative. Most of these early attempts proved unfruitful, but it was two scholars of this period at the University of Oxford (which had been founded by refuges from Paris) who made the key breakthroughs which evolved into modern science. The first of these was Roger Bacon (whom I shall discuss later; his major work was slightly before the 1277 condemnations but can be thought as being as part of the same intellectual movement that led to the condemnations); the second Thomas Bradwardine.

The University of Oxford already had a strong reputation from scholars such as Roger Bacon, William of Ockham and Duns Scotus, when the first of the Merton Calculators, as they came to be known, was elected a fellow at Merton College. Bacon made several important contributions to the emergence of experimental science, as I will mention below, but I think that Ockham and Scotus' contributions were more negative. Scotus emphasised reason greatly, but still felt that Aquinas had gone too far in saying that we could know much of God's nature through reason. God had freedom of action, which meant that physics would not be as regular or worth studying as Aquinas had suggested. Ockham is best known for his Razor, 'Multiple entities should never be invoked unnecessarily', but he generally held and argued, in contrast to Scotus and Aquinas, that we cannot know the world through reason: it

is simply not powerful enough a tool. Instead, we have to rely on observation and revelation. This led the way to the later nominalists and empiricists, however it is dangerous because it undermines the importance of theory and mathematical and geometrical knowledge. At the time, however, it provided a useful counterbalance to the reason-based Aristotelian physics. His solution to the problem I mentioned from Aristotelian physics – why does a projectile keep rising even after we have released it – was that the ball provided its own motive force, which is a similar idea to the theory of impetus. However, his whole discussion of motion was to deconstruct it: he said that it is no more than an object occupying successive places. Movement as a real concept was cut out by his razor.

Thomas Bradwardine was elected to Merton College in 1323, and stayed for a little over a decade before being appointed as Archbishop of Canterbury. He gathered around him a group of scholars who would become among the most prestigious of the medieval era, and who ought to be regarded as among the most prestigious of any era[3]. His primary achievement was that he was the first person to successfully think mathematically about physics. Not just shape and astronomy, but every phenomena of physics: temperature, motion, humidity, strength and brittleness, density and light. Everything he would put a number to, and try to investigate mathematically. He would come up with algebraic formulae linking various quantities. To us, this step might seem obvious: in the fourteenth century, it was a revolution. He said that numbers were a vital ingredient for a successful natural philosophy, that mathematics was the revealer of every genuine truth, and that progress in physics was impossible without it. He wanted to convert what Aristotle said about motion into a mathematical function, and was the first to succeed in writing down a set of equations that summarised Aristotle's physics. Of course, this equation was wrong because Aristotle was wrong – but it was nonetheless the key advance that allowed modern science to be born. He had shown that it was possible to map the laws of motion into an equation. What is of most interest is that his formula made use of logarithms, which officially were not discovered for another two hundred years, and also came to understand exponential growth. Bradwardine also recognised that a heavy body would not fall faster than a lighter body. He was not the first to question this part of Aristotle's thought that was Philoponus of Alexandria, whose works were not known in Western Europe at the time – but it was one of the inspirations behind Galileo's famous experiment.

Bradwardine inspired other scholars at Merton. William Heytesbury, who was the first person to give serious thought to acceleration and its connections to forces, and also was the first to derive the mean speed theorem that an accelerating body moves the same distance as one with a constant velocity at the average speed of the accelerating object. He did not, however, apply this theorem to gravity; that was done by Dominic de Soto shortly before the time of Galileo; but they did establish the underlying mathematics which Galileo later used in his studies of gravity. Richard Swineshead, regarded as the greatest of them all and among the smartest people to have ever lived, wrote the definitive work on medieval mathematics, summarising all that was known thus far and applying it to various physical and imagined situations. His textbook, which included and improved upon Bradwardine's ideas about the construction of a mathematical physics, remained the definitive work on mathematics for the next few centuries, and would have been studied by Galileo. He was

[3]Unfortunately, these early medieval scientists were neglected by the most celebrated thinkers of the renaissance period and, more than that, treated with scorn and derision. Even today many history of science texts start with Copernicus and jump on to Galileo, and attribute all of the achievements of the medievals to Galileo. The achievements of the medieval mathematicians were neglected by many, but not everyone, and they were preserved and admired in certain universities. The most important of these places where the Merton Calculators held particular later influence was the University of Padua, where Galileo later taught and absorbed their achievements. Galileo himself cited his predecessors in earlier works and private notes, so he was certainly aware of and strongly influenced by them. However his ideas were popularised by his dialogues which contained few citations. It is only recently that scholars have started examining the medieval writers in detail, and discovered that many of the ideas commonly attributed to Galileo actually pre-dated him by several centuries. Unfortunately, this has not yet permeated into the public consciousness.

perhaps the first to propose taking simple models, removing most of the complications, to gain an initial general understanding of the most important laws. The calculators also investigated instantaneous velocity, foreshadowing the modern definition of velocity which uses calculus.

The Merton calculators did not perform experiments. What they were interested in was testing out various fictitious circumstances and laws of motion, and trying to compute the consequences of them. In addition to the physics, they gave important advances in various branches of mathematics, such as trigonometry. They provided the right methodology, and the right way of thinking about physics, but not the right answer. Indeed, what would prove to be the key branch of mathematics, calculus, was still unknown to them (they did not develop their thoughts in this subject anything like enough for it to be of use). However, before them it was all but unthinkable to combine mathematics and physics. After them, it was unthinkable not to, and that was their most significant contribution to modern physics. Their school, however, was destroyed when Oxford was ravaged by the black death. It would be some time before Britain would make a similar contribution to mathematical physics.

The work of the Merton calculators was taken up by various continental writers, particularly at the university of Paris. John Buridan (Thijssen and Zupko, 2001; Klima, 2008) was the first of these; in particular he focused on Aristotle's theory of throwing. He was influenced by Ockham, and saw the importance of experimental tests. To him, the idea that the air continually pushed a moving object was ludicrous, because he saw no evidence of it. Instead, he developed the theory of impetus. That is, when we release a heavy rock, we give it impetus which makes it continue in their direction of motion. Falling bodies increase their impetus downwards. He also saw that the impetus must have a magnitude of the body's weight times its velocity. Impetus is not quite the same as momentum (being something external to the object, while momentum is internal to it), but the idea was heading in the right direction. One consequence of his theory was that impetus would continue forever if not checked by an opposing resistance. He also suggested that this might be the cause of the continual and unchanged orbits of the planets, since they, being made of quintessence, were not subject to friction. The impetus of heavenly bodies compelled them to travel with constant angular velocity unless acted on by an external force; that of earthly bodies to travel with constant linear velocity. This distinction between the motions of earthly and heavenly bodies was, of course, disputed by Newton, although, as we shall see, the idea that 'inertial' motion alone could explain the orbits of the planets was resurrected in general relativity. He raised awareness that it was relative motion that we observe, not absolute motion. Buridan also considered the possibility that the earth might be rotating, believing this to be more elegant than a stationary earth and rotating heavens. If we were rotating with the earth, then we would not notice its motion. However, he did not quite solve all the problems that had been listed against the idea that the earth rotated, and ultimately rejected the idea. However, by solving some of the problems, he paved the way for the later mathematical scientists.

Nicolas Oresme followed Buridan, and he carried on the ideas of relative motion and the rotating earth to completion. He recognised that we would not experience a massive wind if the air moved with us. An arrow fired upwards would not fall behind us after the earth rotated and carried us further along, because it would also have the same horizontal impetus that propels us. This challenged Aristotle's law that different types of motion could not happen at the same time. He decided that there was no way we could determine from observation whether or not the world was rotating or moving; so instead turned to the Bible to conclude that it was stationary[4]. Oresme also took up the work of the Merton Calculators, and found that it was possible to reproduce much of their work graphically, plotting speed against time. Other writers of the period moved this on further, for example

[4]Psalm 93:1. Christians today would regard this verse as allegorical, a poetic device; which is quite reasonable given that the Psalms are poetry and abound with metaphors of this nature.

realising that the trajectories of projectiles were parabolas, rather than across and then straight down as implied by Aristotle. Oresme also founded the study of probability as a branch of mathematics, though obviously in a very primitive form.

I should note that these scientists and mathematicians were not Thomists as such. They started by accepting Aristotelian physics, and gradually modified it as it was found to be inadequate, but they did not search specifically for causal explanations. They sat somewhere between the Aristotelian world-view and the modern world-view, accepting (or perhaps proposing would be a better word) a mechanistic model of dynamics, while still maintaining much of the Aristotelian metaphysics and some physical concepts such as the existence of an absolute rest (i.e. absence of locomotion). In physics, they represented a transition between the Greek and early modern perspectives. In philosophy, they would have been sympathetic to some aspects of modern philosophy (more so than their contemporaries), but would have rejected most of it as ill-conceived and lacking the rigour of the developed Aristotelian thinking: they were for the most part convinced disciples of Aristotle or Plato. High medieval Europe, more than any time before or since, was distinguished by its devotion to logical rigour.

Intellectually, Oresme's work on moving bodies was the immediate step before the advancements made by de Soto and then Galileo; his work on graphs the immediate precursor to the algebraic geometry of Descartes; his work on instantaneous motion the immediate precursor to the early studies of calculus by Fermat, Wallis and Barrow. The two hundred and fifty year gap between Oresme and these first of the modern thinkers coincided with the decline of medieval culture following the black death and the wars and schisms that ravaged Europe in the fifteenth century, and then the renaissance, a period when there was only limited progress in either mathematics, philosophy or terrestrial physics, particularly in comparison to what came both before and after it. There are, of course, a few important renaissance scientists, philosophers or mathematicians: the exceptions (perhaps most importantly Copernicus, who was strongly rejected by his contemporaries), stand out because of the general intellectual atrophy of the period. The black death killed the medieval scholarship; the renaissance ensured that it was not revived.

Buridan was regarded by his contemporaries as a truly first class intellect, one of the few thinkers on a par with the great classical philosophers, and with considerable justification. I have focused on his role in encouraging and expanding the work of the Merton calculators, since that was his most important work with regards to my overall thesis, but he wrote profusely on everything: philosophy, theology, law, ethics, mathematics, music, natural philosophy; combining commentaries on the older writers as well as his own original contributions. His writings became the standard textbooks across Europe's universities. They were comprehensive, wide-ranging in scope but detailed in analysis, presented the best overview of both ancient and recent controversies, and were rigorously and in some places brilliantly argued.

They were also immensely dull, boring and difficult to read. Buridan was a brilliant man, but not a brilliant writer. The black death took the best scholars of the high medieval period, the teachers capable of bringing his works to life. The next generations were taught from textbooks they could not understand by teachers who did not understand them. Not surprisingly, there was a backlash, especially when many ancient texts with much greater literary style were rediscovered. Cicero was certainly a highly intelligent man, but he represented a world-view that had been refuted in late antiquity and by the medieval scholars; Buridan's ideas were far more advanced than his. But Cicero was a joy to read (as were the other newly-rediscovered classical scholars), and so people read him, and were entranced. The brilliant work of the high medieval philosophers was largely abandoned and mocked, not because it was found wrong, but because it was found difficult. This, of course, included the legacy of the Merton calculators and their successors such as Oresme, who were lost as a small detail buried in Buridan's textbooks.

The renaissance was a period which somewhat justified the early modern disdain for their predecessors[5]. The most prominent renaissance writers claimed to be the forerunners of a great period of intellectual rebirth; but when we study their actual intellectual achievements we find that there was not really that much of note, certainly in comparison to the periods which proceeded and followed it (though there was progress in art, literary style, engineering and exploration; technological invention was slower than in earlier and later periods, though the printing press remains an important achievement of the period, and the telescope ushered in the modern age). The universities churned out the same syllabus year after year, repeating but not progressing; the scholastic textbooks were still used, but the professors probably did not understand them and certainly were not able to adequately teach them. Philosophy was confused by the popularity of Cicero and stoic thought; where the same terms the scholastics used were given contradictory meanings. Logic, the centrepiece of the scholastic intellectual tradition, was marginalised and would not be revived until the nineteenth century. Mathematics was also marginalised, but fortunately it only required two centuries before the emergence of a new generation of great mathematicians. Intellectual life turned from rational discourse based on logic to a love of symbolism, allegory and imagery; the focus shifted from the sciences to art and rhetoric which tried to depict reality but not truly investigate reality and favoured subjectivism ahead of objective inquiry. Few thinkers distinguished between signs and reality; theology and philosophy declined; the Church lost all moral authority and moral judgement; the slave trade was established: the medievals, who, bringing earlier Christian theology and tradition to its natural conclusion, had abolished slavery in Northern Europe, would have been horrified; renaissance man did not care. It was the age of the inquisition and the witch trials; as superstition replaced rational thought, fear of the occult and the heretic increased and needed to be countered. Alchemy and astrology, long derided by the Church though never quite vanquished, became the fashionable pursuit of what remained of the intellectual class. Of course, by accident, this sowed the seed of the eventual recovery from the renaissance stagnation: though their intellectual foundation was groundless, alchemy provided many of the experimental methods which would develop into modern chemistry and it would inspire the experimental tradition in physics, and astrologers needed accurate maps of the heavens, which at least gave the few genuine astronomers a livelihood. There was throughout Europe an intellectual laziness, coupled with an undeserved arrogance over style, which inhibited progress. This lack of progress was not the fault of the scholasticism which still prevailed in the universities, but rather the renaissance disdain for the medieval scholars and attempt to return to the high point of Rome, and writers such as Cicero who had far greater literary quality but far less intellectual quality than the scholastics.

But the renaissance did, at least, preserve the medieval knowledge and method; it was not a dark age like the centuries between the fall of Rome and the likes of Alfred the Great and Charlemagne. The manuscripts were still there for those who would seek them, the textbooks still available to be taught from, and once the descent into mindless symbolism and representation had run its course, the great men of the early modern period were able to resume science and mathematics where the medievals had left off. But they inherited the renaissance disdain of the Aristotelian philosophical tradition, instead developing the hints of mechanism they found alongside the medieval mathematics into a consistent and complete alternative. Indeed, there is little evidence that they understood the high point of medieval philosophy; certainly they were not able to adequately critique it. In that regard the early modern philosophers were indeed heirs of the renaissance.

But the works of the Merton Calculators, Buridan and Oresme had already spread around Europe, and in a few places they found fertile ground and continued to be taught (as well as Northern Italy, there were some Spanish universities which adopted and preserved

[5]Leonardo di Vinci is the archetypal renaissance man: brilliant artist; brilliant engineer; brilliant visionary; rubbish mathematician; rubbish philosopher; rubbish at foundational physics.

the Mertonian method and made important contributions - of particular importance was Domingo de Soto and his school at Salamanca). In particular, several friars, students from Oxford, returned to their native North Italy and took up positions at Universities such as Bologna and Padua, teaching mathematics, and shaping the curriculum to mirror that at Paris and Oxford. By the end of the fifteenth century, 'Oxford Logic' dominated the North Italian universities. The style of logic was subsequently attacked by the humanists as part of the wider renaissance movement, which increasingly replaced the medieval texts with the ancient Greek works, but Swineshead's and Heytesbury's work continued to be published and in North Italy, and contributed to the University syllabus, particularly at Padua, through to the sixteenth and perhaps early seventeenth century (Lawn, 1993). The legacy of those friars proved to be far more dramatic than could have been imagined, because a few centuries later a brilliant young mathematician named Galileo Galilei[6], who already had an interest in natural philosophy, was appointed Professor of Mathematics first at Pisa and then at Padua, where he was tasked with teaching to that same curricula, and his lectures would have been based on the works of Swineshead and Heytesbury. I do not know how familiar Galileo was with the works of the late medieval mathematicians before his first teaching position at Pisa, but at Pisa and Padua he was certainly forced to master them, absorb them, and be inspired by them in his own work. It was at Padua that he began developing his theories of dynamics, bringing the research of the mathematical natural philosophers to completion.

Aristotle's idea of matter was not questioned so much during the high medieval period. The alternative was atomism, which was seen as being against the doctrine of transubstantiation. While the church was happy to see mechanics developed further, since no key doctrine depended on it, there was more stress when Nicholas of Autrecourt tried to revive the old Greek atomic theory. Eventually he was forced to back down. But his objections found a few ears, and started to spark interest in whether transubstantiation was correct, which when taken up by Wycliffe and became one of the first inspirations of the reformation.

All of these writers were members of the clergy, and some in quite high positions in the church. Thus during the Medieval period, starting from about a century after Aquinas, Aristotle's physics was gradually improved and refined, and the final result was starting to look similar to classical physics. The most important advance of this period was the idea that physics could be treated mathematically. Many of the most important equations describing the simplest types of motion were written down; and Buridan's theory of impetus directly inspired Galileo's conception of inertia, though it was refined by the writers between Buridan and Galileo, and by Galileo himself. The Merton calculators and their successors in Paris essentially solved (either by themselves or by providing all the tools needed for the solution) the problems of dynamics under constant velocity, constant acceleration and constant rotational motion and rotational acceleration (as in pendulums). They provided a much better solution to the problem of projectiles, and proposed ideas concerning terrestrial gravitational force and the rotation of the earth which nearly reached the answer familiar to us from classical mechanics; they could not consider instantaneous changing acceleration because the generalisation requires the mathematics of calculus which they lacked. They defined the concepts of classical mechanics, and their gradual refinement and improvement of the physics they inherited from the Greek and Islamic writers was the key intermediate step between Aristotelian physics and classical physics. They started with something that very closely resembled Aristotle; and they refined it and improved it until they finished with something that very closely resembled Galileo's theory of dynamics, even if they lacked the mathematics to reach Newton. Their biggest weakness, however, was that they had no motivation to test their ideas experimentally.

[6]Also note that, some time earlier, Copernicus had studied at Padua as well; albeit in the medical school.

The geometrisation of physics had an important philosophical consequence: it changed the focus of physics. This proved crucial to the development of modern philosophy, and one of my primary thesis in this work is that, while making physics mathematical was a huge leap forward for physics, it inspired a disastrous wrong turn in philosophy which has endured to the present day. The new breed of geometrical physicists concentrated on what could be expressed in mathematical terms, i.e. that which could be reduced to number or that which could be reduced to shape or otherwise related to the work of Euclid. While motion and locomotion were part of Aristotle's physics, they were not the main focus of the Aristotelian system, which concerned causality. But causality in general is not something that can easily be treated geometrically. Certain types of causality can be considered at least partially in geometrical terms, for example mapping out the paths of moving bodies after a collision or some other such event. But the idea, for example, of final causes, that is that a certain type of body might be predisposed to a certain effect involving a different type of being, is much harder to treat mathematically. Thus the new physics could, due to its methods, only deal with certain phenomena that Aristotle's system could consider, not the entirety of what Aristotle considered. A great deal of Greek natural philosophy was shut out at this time, not because it had been proved that such phenomena were unimportant, but because they simply could not be treated by the new methodology. The early geometrical physicists up to Buridan and Oresme knew and recognised this and carefully avoided the philosophical error; but the generation of Galileo and Descartes forgot this, and started treating the focus of the geometrical physics as though it were the whole of physics. Certainly within that realm the new mathematical approach proved to be far easier and more powerful than the old non-mathematical approach of Aristotle; and its successes soon started pouring in. Aristotelian causality and metaphysics was not refuted by this advance, merely sidelined because it could not be treated using the new method.

But doesn't the success of geometrical physics validate the philosophical idea that natural philosophy can be reduced to that which can be treated geometrically, as Descartes proposed? The difficulty with this objection is in how one measures success; there is an implicit implication in the question that it is restricted to those things which the geometrical approach to physics can study. That mathematical physics is exceptionally powerful in those areas in which it can be applied cannot lead us to the conclusion that there might not also be other aspects of natural philosophy which are beyond its scope and which are better expressed non-mathematically. Nor at this stage do I wish to imply that there are aspects of natural philosophy beyond its scope. It is simply a question which the methodology of geometrical physics alone is not equipped to answer. I do not think that physicists can even properly frame the question until we are approaching a full and complete understanding of the final theory of mathematical physics. Only then can we stand back and ask what are the scopes and limitations of the method, if there are any.

6.5. The Experimental Tradition

Aristotle himself was someone who combined theory with observation, and recognised the importance of both. Unfortunately, this was to a large extent rejected by his successors, with the exceptions of the Astronomers and Philoponus. The Arabs, to their credit, re-established observational investigations, and thus when the Arabic scientific corpus was translated into Latin, there was an example to follow and to be improved on. The first Western European scholar to revive the careful experimental tradition that I am aware of was one Peter the Pilgrim, who wrote a treatise on magnetism to relieve the boredom of besieging a castle in thirteenth century Italy. His small work consisted of careful and systematic observations, many of them quite accurate. His influence, however, was more in terms of the movement he inspired.

It is wrong to say that Aristotle's followers did not value observation as a test and inspiration for their understanding. On the contrary, they emphasised that all knowledge begins

in the senses; they were inspired by and justified their beliefs on the basis of how they saw the world operating; and they were more than willing to change their views when contradicted by observational fact. But there is a difference between observation and experiment. Experimenters are far more meticulous and careful.

(1) An experimenter will keep a careful record of everything he has done; not only the results but the small details of his method. He will publish this in sufficient detail for others to check and try to replicate his findings.
(2) An experimenter will control everything outside the specific thing he is studying, and isolate his experimental equipment from all other influences, so that he can be sure that the result he finds came from the cause he is investigating. If appropriate, he will also set up other studies where he does not make any changes to the set-up, so he can compare the differences between the two experiments.
(3) An experimenter will measure his input and the response, and as much as he can of his equipment, numerically, representing each quantity with a number.
(4) An experimenter will repeat his experiment numerous times, so he can be sure of his findings and reduce statistical uncertainty.
(5) An experimenter will attempt to model his result in terms of a mathematical function that predicts how the system responds to whatever change is made in the input.
(6) An experimenter will keep track and acknowledge any imprecision or uncertainty in his results.

So why did these natural philosophers neglect experiment? There were various reasons, but I will highlight Aristotle's division between natural and artificial motion. When a projectile falls to the ground, it is moving according to its nature. When it is being thrown into the air, its movement, at least for a while, is an artefact, forced upon it from the outside, and contrary to its nature. The purpose of natural philosophy was to uncover the natures of things, and this can only be done by studying their natural motions. An experiment, however, is by nature an artificial setup. You constrain the environment of the object of study, limiting how it can move. Since everything is interconnected, there is no guarantee that constraining its motion in one respect (so you can focus on another aspect of the motion) won't also affect the feature that you are trying to study. So is what you observe a consequence of the being's nature, or of the artificial environment? How can you say what extent your results are a reflection of the nature of the being (which is what you want to understand), and to what extent they merely mirror what you put into the experiment (which would be worthless)? Better just to observe objects in their natural environment, and remove this doubt.

The experimental method, on the other hand, assumes that it is useful to isolate each part or attribute of the being from the rest, and once you understand all the individual parts, you can understand the being as a whole. Thus constraining one aspect of a being still lets it move according to its nature in other respects. This assumption, which in part arose from the growing experimental tradition of the middle ages, formed a key part of the philosophy of science that became dominant in the sixteenth century.

There is some truth to the scholastic objection to experiment, but not enough to mean that we shouldn't conduct experiments. What we are interested in is how the being reacts to certain stimuli. A partially artificial stimuli is just as valid, and tells us just as much about the nature of the being and its reaction to particular forces, as a purely natural one. But the scholastic philosophers were right that experiment alone is not enough to understand the nature of a being. We need to construct a model of that nature, one which can be used to explain both its natural and artificial motions. Studying artificial motion can test that model just as well as non-experimental observation.

Use of the experimental method only developed gradually; as with the mathematical approach to physics it began in the high medieval period, but took centuries to fully mature.

For example, a proper treatment of experimental uncertainty was not possible until the mathematics of probability was sufficiently well advanced, which was not until the nineteenth century. In the early days, the experimental method was applied in individual areas of physics; some started by investigating magnetism, others optics, others the properties of gases, others applied the method to astronomy. Gradually, however, the common thread in all these studies was brought together and the method was improved and refined.

The experimental method has its advantages. It is right to emphasise observation, and to use it to draw out conclusions. It is right to be meticulous and careful in those observations, and to repeat them and carefully record the means of getting the results as well as the actual results so they can be checked and all the assumptions laid bare. However, it also has disadvantages and limitations.

Firstly, as noted above, an experiment is an artificial construct, isolated from the rest of the world. This is necessary to investigate the effects of whatever you are changing; but it is also dangerous, because when placed in an actual situation, the effects you see might be modified by interactions with the environment. This would be fine if physical effects could be reduced to the sum of their parts: one could then study each part individually, and once every component is understood, one would gain an understanding of the whole. Many things are just the sum of their parts (for example, the great mechanical engines of the nineteenth century), and for these the method works well. For others, it works to a reasonable approximation, where the interactions between the parts are slight, so one can start with the findings of the method of examining each part, and then refine it. But there are also strongly interacting systems, where a study of the parts individually does not give us any understanding of the whole. Initial success in areas where it is applicable led people to hope, without strong reason, that everything might ultimately be reduced to its parts; but this is hope and prejudice, not reasoned conclusion.

Secondly, the experimental method focuses by its very nature on observable properties and correlations. This gives a limited view of their nature. It therefore works well if we assume that substances can be defined by their properties; but if this is not true, then it will only work up to a certain point, and not allow us to gain any further understanding. Whether substances are more than just their properties is not something that the experimental method can address by itself. Equally, the experimental method gives us no understanding of universals; it is not capable of affirming or denying them, it simply cannot cope with them.

Thirdly (and related), by attempting to model and fit its results using mathematical functions, it is not creating a theory or generating understanding. It is simply a study of correlation; it is an initial guess. It might have some fundamental importance, or it might not: at this stage, we simply cannot say. A theory starts from first principles, and attempts to derive the experimental result (or something else which equally well explains the data) from those principles. The modelling done by an experimenter gives no understanding; only theory and mathematical modelling can allow us to understand (at least partially) why the object of study behaves in that way in terms of more fundamental principles, which is the ultimate goal of science. The experimenter can only describe how it behaves, and even there without any degree of certainty, due to the limitations of experimental imprecision and the enormous dangers of extrapolation or statistical regression without a guiding theory. However, the experimental model can provide a useful inspiration for the theorists, as well as providing a test of the theoretical calculations.

Now these limitations of the experimental method can be overcome by combining it with the mathematical approach. The mathematical approach to physics also has its weaknesses, but they are, for the most part, different weaknesses. By using experiment and mathematical theory together, we can formulate an approach which avoids all these drawbacks. And this, of course, is what scientists, particularly physicists, do.

However, there are weakness that both the experimental and mathematical methods share. Most importantly, they both try to express their objects of study in terms of numbers. This is very good, for those things which can be expressed numerically, such as lengths or velocities or accelerations. But many things cannot be whittled down to a sequence of numbers; and for these concepts and objects of study neither the experimental method nor the mathematical method nor their combination can offer us any guidance. This is not a conclusion of the scientific approach; not even a premise, but only a methodological choice. Not everything can be expressed as a number; and for those things which cannot be expressed numerically we should not even attempt to try, because any attempt will just mislead us. We will end up measuring something different from what we want to be measuring, or think that we are measuring, and thus come to conclusions which are wholly wrong. That doesn't mean that we cannot study such things, but we must find a different method. Thus the experimental method is limited in its scope; and, unfortunately, the limitation is the same as in the mathematical approach to physics. We should not assume from the success of modern science that everything can be quantified; we can only suppose that everything which science has successfully studied can be quantified. Success is measured by how well something achieves its purpose; the success of science means that it does a good job of understanding the aspects of reality that can be quantified. It says nothing about those aspects of reality which cannot be quantified, but it is not the purpose of modern science to do so. To consider such things (even if it is just to deny their existence), we need to turn to a different method of gaining knowledge. We should not, therefore, come to such a philosophical conclusion that if it isn't scientific it isn't real on such shaky grounds.

It is worth listing out these methodological assumptions (which are in some cases opposed to the premises of the Aristotelian metaphysics), because people were deceived by them and the initial success of the experimental sciences into thinking that they are genuine features of reality. However, this is not reasonable: that something works well in one area does not mean that it will work well in all areas; particularly when 'works well' is redefined to mean only gaining a shallow understanding of those things which the scientific method studies, and not a complete understanding of their natures, since that is only as far as the experimental method alone can take us.

Experiment was key in one of the pseudo-sciences that became popular during and after the medieval period, namely alchemy. The alchemists did not succeed in turning base metals into gold. Indeed, they usually ended up with less gold than they started: it was an expensive pursuit. But they did make significant progress in various areas of chemistry. For example, they discovered acids and their uses, and extracted many new types of metals. But what they valued most of all was careful, meticulous, experimentation.

The scholastics, or certainly the best of the scholastics, strongly believed that observation was the test of scientific knowledge. After all, Copernicus' theory was rejected on empirical, not philosophical, grounds: it made certain predictions which at that time were contrary to the best observed evidence. Similarly, Aristotle's dynamics was accepted because it worked very well as a description of the motion of bodies in a fluid such as air or water – which is where people made their observations. The Aristotelian scholars frequently backed up their speculation with allusions taken from observed nature. They fell short not because they were unaware of the importance of observation, but because they were not careful and systematic enough in making their measurements – this is the difference between observation and experiment. Those who observe see the arrow leave the archer in a straight line, and see it fall down on the opposing army almost vertically (because by then air resistance has killed almost all of its horizontal momenta), and don't question Aristotle – who stated that it was given an artificial motion by the archer which eventually slows down as the arrow returns to its natural state of rest, and then it switches abruptly to its natural motion of descent. An experimenter, on the other hand, carefully plots the position and velocity of the arrow at each point along its trajectory, and discovers that the initial motion

is not quite linear, but instead approximately parabolic, and that the final descent is not quite vertical, and the velocity changed smoothly. But obviously, with medieval technology, the speed of an arrow, and the difficulty of performing impassioned science when people are shooting arrows at you, this sort of detailed measurement is almost impossible. Aristotle's physics worked within the experimental errors available at the time, and on that basis, given that there were no reasonable alternatives, it was accepted. As the experiments improved, and problems were found in Aristotle's physics, it was gradually rejected, first in favour of the theory of impetus, and then the theory of inertia, and finally Newton's three laws.

Indeed, the empirical tradition arose in the late medieval period in part as an attempt to go beyond what Aristotle wrote. Why should God be constrained by what a one and a half thousand year old (as Aristotle was in the time period when this questioning started) pagan Greek philosopher wrote? Why should He be forced to order nature in that way? But to go beyond Aristotle, the early scientists needed to appeal to an authority greater than Aristotle, and there were only two easily available: the geometry as adopted by the Merton Calculators and their successors, built on the 'unquestionable' foundations of Euclid's axioms, and secondly observation of the natural world. The empirical method was seen as a way of discovering more about God by study of His works.

The main difference between modern science and the earlier tradition was not then in an understanding of the importance of observation, but in an improvement in the method of observation, and in particular a growing understanding of the importance of understanding experimental error and of setting up the experiment so that the errors were reduced (of course, I am being anachronistic in the language I use here). This improvement in observational methodology was a gradual process which took place over the course of the middle ages, reaching a climax with Galileo, but continued to be developed as the scientific age progressed. The flaw of the scholastics (I should say the worst of the scholastics; it was not true for all of them) was that they believed that naked eye observation was enough.

One of the early alchemists was Roger Bacon, who, after losing all his money, became a friar and moved to Oxford University. His predecessor at the university, Grosseteste, had emphasised how important experience was to Aristotle. Bacon took that further, and added his own experiences into his investigations into natural philosophy. Unlike most academics of both the classical and medieval period, he had a respect for craftsmen. He was not afraid to get his hands dirty. He is widely regarded as the founder of the modern experimental sciences. He is also widely regarded as one of the first scientists to get into trouble with the church, although this tale is based on late sources, and might not be accurate.

Roger Bacon is well known for being an early advocate of the method of induction, and that these conclusions should be subsequently tested: a method of verification. He strongly emphasised, against many of his contemporaries, that facts should be gathered first, ahead of speculation. His method was observation, hypothesis, experimentation, and verification. He left behind detailed experimental notes, for the benefit of other scholars. He was critical of the minute and detailed speculations of the scholastic philosophers, who would argue for ages about arcane pieces of logic. Bacon was not an isolated scholar; many of his views had been previously aired by others including Grosseteste, but he was an influential one. His most important contributions to experimental science were in optics, where he contributed to the discovery of the lens. However, he never fully solved the problem of the eye, and died shortly before lenses were produced in sufficient quality to give the magnifier, spectacles and eventually the microscope and telescope. He was also one of the first Europeans to describe the manufacture and importance of gunpowder. Although known today as one of the first to suggest the modern scientific method, he was primarily a scholastic philosopher, and not so distant from his contemporaries as is sometimes claimed.

Experiment was seen as important subsequently to Bacon, although until Galileo it was more prominent in astronomy, chemistry and biology ahead of physics. It was in this period that the first Western (the Chinese had got there earlier) detailed studies of anatomy were

carried out. There were also important empirical studies of magnetism and optics, and later on gases; the experimental tradition thus continued in the background through the works of numerous minor scholars (whose most important contribution to science was just to keep the empirical method alive) until Galileo and Newton fully merged it with the geometrical tradition.

The fifteenth century writer Nicolas of Cusa also made an important set of contributions. He is known for two texts, firstly *On Learned Ignorance*, where he hypothesised about such crazy notions as that the earth was not the centre of the universe, but rotated around the sun, and that the sun was just another star and the earth just another planet. In his work *On Measuring*, he was perhaps the first to set out in full the method of empirical investigation. He recognised that rough observations were not enough to merge mathematics with observation, but that one had to be careful enumerating the observed quantities such as weights, times and distances. Though this painstaking and careful research was difficult, it was necessary to properly quantify observation.

Probably the man most associated with emphasising the experimental method was Francis Bacon, who was active right at the start of the seventeenth century. He was primarily a statesman, working in the court of James I of England (and VI of Scotland), but also produced various philosophical works, and was renowned as one of the brightest men of his period; certainly he was one of the most influential. Not because of any experiments he did – his contribution to science was only in terms of methodology. It is important to note the time period. This was, of course, at the height of the renaissance, so the emerging intellectual atmosphere was very much opposed to the medieval philosophy; and in this sense Bacon was very much a man of his times and culture. However, he was also raised in firmly Protestant England after the Reformation. Britain was now the foremost Protestant power. It had just beaten off the best the Roman Catholics had thrown at it militarily, but there was also an intellectual battle. The Anglican theology was, like all the early magisterial reformers, still greatly influenced by Scholasticism: their aim was to reform the philosophical basis of Catholicism, removing what was wrong and keeping what was good. Hooker, for example, the leading English theologian of the period, argued against the Roman Catholic excesses, but was clearly still operating from the same basic intellectual world-view, using the same definitions and adopting many of the same premises. However, other streams of Protestantism were against everything Catholic, and scholastic philosophy was very much a Roman Catholic endeavour, and at the heart of the rational basis of the Roman church. It was easy to see how opposition to Catholicism led to a general prejudice against Catholic philosophy. That is not to say that Bacon himself was guided by prejudice – he was a highly intelligent man, and motivated by the flaws he saw in the Catholic methodology; his arguments need to be answered or accepted rather than dismissed because of the context in which they arose. But it is important to bear in mind the culture in which he operated and was received.

He emphasised the method of inductive knowledge. He believed that the way to understand things was to first gather as much data as possible, and then draw conclusions from it. Unlike his namesake three centuries earlier, he did not combine an understanding of the importance of experiment with an appreciation of classical philosophy. On the contrary, he was entirely distrustful of the power of human reason, trusting only his senses. He was thus in many respects one of the founders of the philosophical school of empiricism.

His most important work was the *Novum Organum* (originally written in Latin), named after a treatise by Aristotle on Logic, in which he argued against the Aristotelian deductive method and in favour of a wholly inductive method. My intention is to briefly review certain parts of this work where he firstly sets out his method, and secondly the portions where he responds to Aristotle's metaphysics. I am highlighting Bacon because he was one of the most influential figures of his era, one of the clearest advocates of the experimental science, and one of the most vigorous opponents of Aristotle and the Medieval consensus. He is

both a typical and outstanding representative of the anti-Aristotelian school of his period, and his contemporaries (at least among the other early empiricists) did not offer any better arguments than his own. The first part of his work outlines his general method; the bulk of the second gives his own thoughts on the scientific ideas which could be deduced from his method, and this is less interesting because his proposed physics was generally wrong.

He was a bitter opponent both of the Medieval/Aristotelian consensus, and those such as Descartes who (albeit a generation after Bacon) adopted a sceptical approach, as seen in his preface.

> Those who have taken it on themselves to lay down the law of nature as something that has already been discovered and understood, whether they have spoken in simple confidence or in a spirit of professional posturing, have done great harm to philosophy and the sciences. As well as succeeding in producing beliefs in people, they have been effective in squashing and stopping inquiry; and the harm they have done by spoiling and putting an end to other men's efforts outweighs any good their own efforts have brought. Some people on the other hand have gone the opposite way, asserting that absolutely nothing can be known – having reached this opinion through dislike of the ancient sophists, or through uncertainty and fluctuation of mind, or even through being crammed with some doctrine or other. They have certainly advanced respectable reasons for their view; but zeal and posturing have carried them much too far: they haven't started from true premises or ended at the right conclusion. (Bacon, 1604)

His method was to gradually increase the amount of certainty. He criticised the Aristotelians for jumping straight from sense data to what they believed to be the most general foundational principles, and then proceeding by deduction from these principles. This he believed to be contaminated by various sources, in particular cultural prejudices and the natural presumptions of man's intellect. He believed that we are overly keen to make abstractions, and are overly biased by what we would like to be true, making unaided reason next to useless. He asserted that these methods had made no progress whatsoever. Of course, if the premises were false, then so would be the conclusions of the deduction. Instead he advocated a method of starting with what was accessible to the senses, and try to understand the basis of those sensations, then to move to the layer that underlay that, then the layer that underlay that, and so on until eventually this reached the foundations of scientific knowledge. This, of course, presupposes that scientific knowledge does consist of layers which can be peeled off one at a time (though to be fair to him, this is not too far from what a modern physicist would say). He set himself as the opponent of the Greek and medieval scientists more in terms of method than anything else; he also disagreed with their conclusions, but this was not where he directed his disputation.

His work was divided into numerous numbered short statements or aphorisms, the first of which was

> 1.1. Man, being nature's servant and interpreter, is limited in what he can do and understand by what he has observed of the course of nature - directly observing it or inferring things from what he has observed. Beyond that he doesn't know anything and can't do anything. (Bacon, 1604)

He followed this up with a fierce criticism of previous work:

> 1.5. The mechanic, the mathematician, the physician, the alchemist and the magician have all rubbed up against nature in their activities; but so far they haven't tried hard and haven't achieved much.

> 1.6. If something has never yet been done, it would be absurd and self-contradictory to expect to achieve it other than through means that have never yet been tried. (Bacon, 1604)

This argument, repeated in various forms throughout his work, was at the heart of his objections to Aristotelianism: it has not made any progress, so the methods should be abandoned as useless. Note the targets of his criticism: the physician refers to those who follow Aristotle's physics; the mathematician refers to the Merton calculators and their successors, who included Bacon's near contemporary Galileo; the mechanic refers to those who adopted the mechanical world view, which inspired Descartes. In short, Bacon's criticism included those who were achieving great things even as he was writing, or in the next few decades. Of course, his statement that it had not made any progress needs to be examined against the evidence, and it is found wanting. The medievals were making progress, certainly up to the time of the black death, and if they had stalled after that it was more due to the ravages of disease and war and their effects on society than anything else. We have already seen that the mathematicians had already solved many of the limitations of Aristotelian mechanics, and, of course, Copernicus had already put forward his revolutionary hypothesis and Brahe and Kepler were contemporaries of Bacon and already independently working on their improvements to Copernicus' model. Indeed, the the most important missing ingredient needed to reach classical mechanics was not observational, but mathematical: calculus. Thus the statement that the earlier thinkers had not achieved much was ill-timed: they had almost got there. However, as with most statements, there was a kernel of truth in what Bacon had said: Buridan and Oresme's work needed to be tested by experiment, and that would have to wait for Galileo.

And this is my main criticism of Bacon: he was right to realise that careful and controlled experiments were necessary, but they were needed as an addition to the older methods, not as the replacement that he advocated. It was through the combination of methods that progress was made, not through one approach alone. As he said the mathematicians had not yet succeeded; but it was going to be a mathematician who would succeed. The issue was not that they tried mathematics too much, but that they did not yet know enough mathematics.

Another point about Bacon was that he was keen on practical uses of natural philosophy: new inventions were the guide to scientific progress. He thus defined progress solely by progress in the experimental sciences. In other words, he judged philosophy by how well it explained the observational world. Not it is clear, however, this can only lead to scientific knowledge. Bacon by definition excluded the less obviously practical philosophies from his measure for success. This, however is a bad idea. Each field of knowledge should be judged by the standards appropriate to that field. Science indeed should be judged by how well its predictions match up with experiment; but the measures of success for mathematics, metaphysics and other fields are different because the topics of study of these fields are different.

I should also highlight the following passages, which are typical of the types of mistakes that Bacon made:

> 1.14. A syllogism consists of propositions, which consist of words, which are stand-ins for notions. So the root of the trouble is this: If the notions are confused, having been sloppily abstracted from the facts, nothing that is built on them can be firm. So our only hope lies in true induction.
> 1.15. There is no soundness in our notions, whether in logic or in natural science. These are not sound notions: substance, quality, acting, undergoing, being; And these are even less sound: heavy, light, dense, rare, moist, dry, generation, corruption, attraction, repulsion, element, matter, form and so on; all of those are fantastical and ill-defined. (Bacon, 1604)

The first of these is an example of the false dichotomy, which Bacon seemed to use continuously. He is right to say that a syllogism built from sloppy premises is infirm. However, he then jumps to induction as the only hope, as though deduction from false premises and induction were mutually exclusive. On the contrary, the best approach turned out to be deduction with experimental checks of the conclusions in order to test the premises. Bacon did not seem to realise that there were more options than pure deduction and 'true' induction.

The second point is where he argues from assertion, that there is no soundness in our notions. He offers no evidence for this statement. He does not engage with why these concepts were first accepted, and show that reasoning to be false. His criticism of the basic scholastic notions was that they were ill-defined. If true, that would mean that rather than immediately abandoning the concepts, we should first seek to define them more carefully. A definition, of course, need not be perfectly precise, and certainly cannot be at first while people are still limited in their understanding, but this does not make the concept unsound. It is better to be vague and right than precise and wrong; and until very recently (and perhaps not even now) it has been impossible for us to be precise enough and confident that the uncertainty is under control; if Bacon's predecessors were vague, then at that time it was right for them to be so because they were aware that they did not know enough to state things precisely.

Bacon establishes his method with another false dichotomy:

> 1.19. There are and can be only two ways of searching into and discovering truth. (1) One of them starts with the senses and particular events and swoops straight up from them the most general axioms; on the basis of these, taken unshakably true principles, it proceeds to judgement and the discovery of intermediate axioms. This is the way that people follow now. (2) The other derives axioms from the senses and particular events in a gradual and unbroken ascent, going through the intermediate axioms and arriving finally at the most general axioms. This is the true way, but no-one has tried it.
>
> 1.22. Both ways set out from the senses and particular events, and come to rest in the most general propositions; yet they are enormously different. For one of them (1) merely glances in passing at experiments and particular events, whereas the other (2) stays among them and examines them with proper respect. One (1) proceeds immediately to laying down certain abstract and useless generalities, whereas the other (2) rises by step by step to what is truly better known by nature. (Bacon, 1604)

This is false, firstly because it excludes the method of using experiment to test propositions, but also because it sets his two methods against each other. Rather, the best approach is to use all three of these methods together, letting insights from one influence how we proceed with another. Indeed, why should we assert that the general axioms derived from Bacon's method (1) are necessarily useless? Certainly, they might sometimes be useless, but not always. Certainly we see general principles before we go into the details of any specific experiment: that the universe is regular and experiments repeatable; that there is both change and constancy; that there are many different ways in which something can be the same thing; that the same matter can make numerous different types of objects, and thus matter alone cannot be the sole principle behind material objects. These ideas, universally seen, lead to the generalities of metaphysics. They provide (according to the Aristotelian) the presuppositions which are needed as a prerequisite to starting empirical investigation. Like everything else, the experimental method is built on various premises, and these premises cannot be established by experiment. Bacon's method (2) if used alone is impossible: one has to make certain general assumptions before one can begin careful empirical investigation. What Bacon did was to normalise those assumptions, assume that

they were self-evident and subsequently assert that he was not making any general assumptions. Here is an example of Bacon asserting as true that which he assumes but does not and cannot prove from his method:

> It's true that nature really contains only individual bodies, performing individual pure actions according to a fixed law; but in science this law is what we inquire into, discover, and explain; it is at the root of our theorising as well as of our practical applications. (Bacon, 1604)

This is just as much a presupposition as the Aristotelian system; it is not something we determine by Bacon's method. A presupposition, of course, is required, because before you begin a scientific search it is necessary to know what sort of thing you are looking at. It is therefore necessary to presuppose some metaphysical principles before one starts in physics. Bacon, like many contemporary readers, perhaps failed to notice that they were doing that with this thought because the mechanistic system is so absorbed into our way of thinking. But that doesn't mean that it isn't a presumption made by a glance in passing at experiments and particular events. This argument that Aristotle's was inferior to his system thus fails because even if the accusation is true, Bacon was also guilty of it. Of course, one should judge between these different sets of presumptions based on which best fits with the theories tested by experimental evidence (if there was any difference determinable by experiment), and I'm sure that Bacon would have sympathised with that view, had he realized that he was making the same mistake; but he did not apply it when examining the medieval system. He dismissed that without any empirical justification.

Bacon did not have a good understanding of the scholastic philosophy; many of the statements he made about it showed naivety.

> 1.48. But this inability interferes even worse in the discovery of causes, and here is how:
>
> The most general principles in nature have to be brute facts, just as they are discovered, and can't be derived from any still more general or basic cause. Yet the restless human intellect still looks for something to explain why they are true. Then in that doomed struggle for something further off, it finds itself defeated, and instead falls back on something that is nearer at hand, namely on final causes – i.e. on the notion of what a principle is for, what purpose explains its being true. Science has been enormously messed up by this appeal to final causes, which obviously come from the nature of man rather than from the nature of the world – that is, which project the scientist's own purposes onto the world rather than finding purposes in it. To look for causes of the most general principles is to do science in an ignorant and frivolous way just as much as not looking for causes of subordinate and less general truths. (Bacon, 1604)

First of all, the most general principles don't have to be brute facts, at least not as I have defined the term (definition 3.2.4). In Aquinas's system, the most general principle is God, who is not a brute fact, since, as a being of pure actuality, cannot be other than what He is.

In this passage, Bacon attacks final causality. Notice how Bacon has misunderstood this concept. He regards it as referring to purposes, rather than, as more accurately reflects the best of the scholastics, merely a list of the possible changes that can arise from a given substance. Secondly, he makes the mistake of supposing that because scientists had inaccurately identified final causes that made the whole concept of final causality incoherent. It doesn't, it just makes some applications of the concept incoherent. The idea of final causality (not design) was advocated as the only way in which the idea that the universe could be regular could be explained. Whether individual final causes have been correctly identified is a separate question. Furthermore, that the final causes have been inaccurately identified in the past does not necessarily mean that they will also be so in the future.

Finally, Bacon asserts that science has been messed up by the appeal to final causes, which *obviously* arise from the nature of man rather than the nature of the world without any evidence to back up his claim.

Next Bacon turns to the concept of forms:

> 1.51. The human intellect is inherently prone to make abstractions, and it feigns an unchanging essence for things that are in flux. But better than abstracting from nature is dissecting it; which is what Democritus and his followers did, getting deeper into nature than anyone since. What we should be attending to is matter, its micro-structures and changes of micro-structure, and actus purus, and the laws of action or motion. The alternative to studying matter is to study forms, but forms are fabrications of the human mind, unless you want to call the laws of action 'forms'. (Bacon, 1604)

Again, note the false dichotomy in the first sentence. Aristotle's system was designed to combine the ideas of an unchanging essence and being in flux; saying that the overall form remained the same while the system moved from one potentia to another. In other words, it is possible for aspects of something to remain constant while other aspects of it are changing; and this is the whole basis of Aristotle's metaphysics. However, Bacon and his successors reverted back to an either/or situation. Bacon makes a further false dichotomy when he asserts that the study of matter and of forms are alternatives. He sets up study of matter against study of form, rather than, as we should, considering them both together. Bacon's assertion that forms are fabrications of the human mind [i.e. don't correspond to something in nature] is an assertion given without any evidence; indeed contrary to the evidence. It is likely that Bacon did not understand the scholastic notion of form; when discussing the micro-structures and change of micro-structure of matter he was discussing part of what goes into the idea of form since the possible internal structures are part of what the scholastics meant by form. Equally, the laws of action and motion are related, summarised within final causality in the scholastic system. Rather than arguing against scholasticism, Bacon was arguing for it, but giving its notions different names to make it appear, including to himself, that he was opposing it. Indeed, later on he contrasted those who dwelt excessively on the particles, and those who dealt excessively on the structure, and recommended that a middle course should be drawn between these two. By emphasising both matter (particles) and form (containing the structure), this middle course is precisely what the scholastics proposed.

Now he attacks Aristotle in general.

> 1.62. ... What has gone wrong in philosophy is that it has attended in great detail to a few things, or skimpily to a great many things; either way, it is based on too narrow a foundation of experiment and natural history, and decides on the authority of too few cases. (1) Philosophers of the reasoning school snatch up from experience a variety of common kinds of event, without making sure they are getting them right and without carefully examining and weighing them; and then they let meditation and brain-work do all the rest. (2) Another class of philosophers have carefully and accurately studied a few experiments, and have then boldly drawn whole philosophies from them, making all other facts fit in by wildly contorting them. (3) Yet a third class consists of those who are led by their faith and veneration to mix their philosophy with theology and stuff handed down across the centuries. Some of these have been so foolish and empty-headed as to have wandered off looking for knowledge among spirits and ghosts. So there are the triplets born of error and false philosophy: philosophies that are (1) sophistical, (2) empirical, and (3) superstitious.

> 63. The most conspicuous example of (1) the first class was Aristotle, whose argumentative methods spoiled natural philosophy. He
> - made the world out of categories;
> - put the human soul, the noblest of substances, into a class based on words of the second intention;
> - handled the issues about density and rarity (which have to do with how much space a body takes up) in terms of the feeble distinction between what does happen and what could happen;
> - said that each individual body has one proper motion, and that if it moves in any other way this must be the result of an external cause,
>
> and imposed countless other arbitrary restrictions on the nature of things. He was always less concerned about the inner truth of things than he was about providing answers to questions – saying something definite. ... Whereas in Aristotle's physics you hear hardly anything but the sounds of logical argument – involving logical ideas that he reworked, in a realist rather than a nominalist manner, under the imposing name of 'metaphysics'. Don't be swayed by his frequent mentions of experiments in his On Animals, his Problems, and others of his treatises. For he did not consult experience, as he should have done, on the way to his decisions and first principles; rather, he first decided what his position would be, and then brought in experience, twisting it to fit his views and making it captive. So on this count Aristotle is even more to blame than his modern followers, the scholastics, who have abandoned experience altogether. (Bacon, 1604)

Once again, there is a lot of assertion here without either evidence or even proper understanding. He presupposes that nominalism is better than realism without evidence; I have already discussed that when measured against modern physics nominalism should be considered untenable. He dismisses Aristotle's categories and theories of motion; and even if Aristotle did not understand enough to correctly categorise or to identify proper motion, that doesn't mean that the concepts are flawed. He makes presumptions about whether Aristotle fit his philosophy to his experiments or his experiments to his philosophy. What he rejected from Aristotle was based only on his own prejudice. His criticism of logical ideas and answering questions seems really odd; perhaps he is saying that these things alone are not enough. If you read a modern textbook on physics, it would be full of mathematical logic rather than syllogisms, but nonetheless Bacon might well reach the same conclusion. Of course, the modern textbook is built on experiment, but that is often left in the background, just as Aristotle left his experiments in the background (although sometimes making reference to them, as Bacon himself noted).

Against the scholastics, if his criticism is that they did not conduct enough or careful enough experiments, then I would fully agree with it. While, more than most modern physicists, I respect the medievals, my admiration does not stop me from recognising their many flaws, nor that their system needs to be revised in the light of modern physics. However, his criticism was rather that they had abandoned experience altogether: this criticism is false, as, for example, their experiments with optics, magnetism and ballistics shows.

> 1.71. The sciences that we have come mostly from the Greeks. For the additions by Roman, Arabic and later writers are neither plentiful nor important, and such as they are they have been built on the foundation of Greek discoveries. Now, the wisdom of the Greeks was that of teachers of rhetoric, and it spawned disputations, which made it the worst kind of inquiry for finding the truth. ... But although the two groups of philosophers were in other ways unalike, they had one thing in common: both lots were teachers of rhetoric; both turned everything into a matter for

> disputations, and created sects that they defended against heresies. They turned it all into 'the talk of idle old men to ignorant youths'. But the earlier of the Greek philosophers ... did not open schools, as far as we know. What they did was to apply themselves to the discovery of truth. 1.78. ... We can properly count only three periods when learning flourished, and they lasted barely two centuries each: that of the Greeks, the second of the Romans, and the last among us – the nations of western Europe[7]. The intervening ages of the world were not flourishing or fertile for the growth of knowledge. (Don't cite the Arabs or the schoolmen as counter-examples to that; for they spent the intervening times not adding to the weightiness of the sciences but crushing them with the weight of their books!). (Bacon, 1604)

I have cited this because once again we now understand well that he badly underestimated the vital contributions of both the Arabs and schoolmen. But equally, we have the implication that the great Greek philosopher's did not apply themselves to the discovery of truth. This statement is staggering; most of the writers might have got things wrong, but it was not because of a lack of desire for knowing truth.

I could add more about Bacon, for example his confusion of being with that being's properties; his removal of formal and final causes from his system without understanding either, his demeaning of even the material and efficient causes that he retained.

I should also say that my opinion of Bacon is not all critical. I have selected passages which I believe are good examples of the sort of mistakes that were made by Bacon and his contemporaries, because it is these passages which are most relevant for my goal of understanding how the modern world rejected Aristotle, but nonetheless he also did get a great deal right. Most importantly, his emphasis on experiment. He correctly emphasises and describes how errors can be introduced and propagated into world-views, and also in the ways in which we should be cautious when using induction. In his emphasis on induction he did offer a corrective on the scholastic and mathematical school which went too far in the opposite direction. He warns against both the excess of being too quick to jump to conclusions, and of despairing of ever finding conclusions. He was also clear about the experimental method: simple experience was not enough, it needed to be done systematically and carefully controlled (for example to evade selection bias). A lot of his criticism concerning how progress is stunted hits the mark, even when applied to today's environment.

I have highlighted Bacon in detail because he is a good example of the renaissance man, with both his advantages and flaws. Our modern views are largely sympathetic with his; it was because of criticism of men of Bacon's era that the modern man, until the latter part of the twentieth century, despaired of the 'Dark Ages', and even now the conclusions of modern scholarship that they were not as dark as we were led to believe has not filtered down into popular consciousness.

But I had to cite extensively from Bacon, as a typical critic of this period, because my main point does not concern what is present in his criticism but what is absent. He does not present any experimental evidence against scholastic philosophy. He does not attack the true premises that motivated Aristotle. He does not go into the details of the reasoning, and show there to be logical inconsistencies. At best he claims that certain key concepts are ill-defined, without justification that they were ill-defined to a dangerous extent, or what constitutes the boundary between dangerous imprecision and dangerous over-precision. When he did speak of scholastic philosophy, he generally got it wrong; and certainly did not discuss any of the details. So without experimental evidence, arguments against the premises, arguments against the details of the logic, or even understanding what he was criticising, we are expected to believe that Bacon offered a devastating critique

[7]He here means the renaissance period.

of the scholastic system. The scholastic philosophical system (as opposed to its notions about physics) was not dismissed for valid reasons. It was firstly misunderstood (with final causality treated as purpose), and then simply abandoned without any direct evidence or logical argumentation. That one explanation works well is no evidence that its rivals are wrong; they have to be shown to be false. With regards to the causal picture of scholasticism, this was not done at the time of the renaissance, and has to my knowledge not been done since. Of course, this doesn't mean that the causal picture is right, merely that it was not shown to be wrong. This is typical of most of the renaissance thinkers: they offered no reasons against final causality; just pushed purpose aside as redundant and ineffective. We must decide between the renaissance preference for mechanism and the scholastic preference for causal explanation (as well as any alternative ideas) by the only tools we have: logic, experiment, and comparison against our best scientific theories.

Of course, the experimental method soon prospered, with Gilbert with his discussion of magnetism, and Boyle with his numerous works in both physics (such as his studies of gases) and chemistry (of which he was the founder) prominent examples. However, it should be noted that neither of these were fully devoted to Bacon's philosophy of nature; Boyle for example still discussed the properties of matter such as heat in terms of the dispositions of the constituent corpuscles, which largely expresses Aristotelian rather than mechanistic notions in a different language (Burtt, 1931). He sat somewhere between the emerging mechanists and the medieval causal physicists. He did not use purposes, but neither did he reject them (unlike Hobbes, Bacon and Descartes): he was just more interested in the question of *how* rather than *why*. (Recall the difference between final causes and purposes: there is nothing in this position which is contrary to Aristotle as I interpreted him in the previous chapter.) But Boyle was influenced by the success of Galileo; he was halfway between the mechanists and the causal physicists, sometimes favouring one, and at other times favouring the other. The purpose physicists still had a few successes: Harvey's studies of anatomy, which were instrumental in the launch of modern biology, were motivated by design principles. But final causality – interpreted as purpose – was being pushed out of physics and into biology, where it would remain until Darwin came along.

The almost immediate triumph of Newton seemed to settle the matter in the best way possible. He showed, at least as far as mechanics and locomotion were concerned, physical processes could be expressed as a general mathematical law, valid both on earth and in the heavens, which treated each bit of matter the same. Over the next two centuries, almost every other known physical phenomena was explained in terms of Newton's original work. Scholastic causal explanations were out; renaissance mechanical explanations were firmly in.

6.6. The Emergence of Mechanism

Mechanism is a metaphysical position, which stands as an alternative to empiricism and the Aristotelian system (among others). It is solely a theory concerning the basis behind physics, and thus is not a complete metaphysics, but instead can be absorbed as part of various fuller philosophies. It provided the inspiration for many advances in physics, to a lesser extent in the high medieval period and then fully from Galileo, Newton and most of their successors until the twentieth century. As such, it is often mistakenly believed to be a immovable part of modern physics; however this is false. It is not directly concerned with the actual specific structure of physical law, but instead tries to describe what physical law represents. For example, Newton's three laws are an example of a physical theory consistent with the mechanical philosophy; but they are not the only physical theory which we could write down consistent with that philosophy. We could consider a system where force was proportional to velocity, or perhaps the rate of change of acceleration, rather than acceleration as in Newton's second law. Both of these are possible mechanistic theories, they just happen to be wrong. If physics, then, is the description of beings as they are,

then mechanism is more general than this, and thus belongs in the realms of metaphysics. Physics is the set of experimental results and the mathematical theorem that connects those results; mechanism is an attempt to explain why there is such a theorem in the first place, and consequently to restrict what sort of theories could mirror reality.

The mechanical philosophy (as it reached maturity in the late renaissance era) can be defined in terms of a number of basic postulates, so a mechanical philosophy will accept a clear majority (though not necessarily all) of the following:

(1) That all compound objects can be reduced to the sum of their parts, eventually reaching the most fundamental objects (which may be fundamental particles in some schools of mechanism, or geometrical objects in others, or extension itself in another school). These parts are not in any way dependent on each other, nor are their arrangements suggested by their internal natures, and do not influence each other except as described by the laws of physics, mostly mediated through collisions.
(2) That the effects of these parts can be determined solely by their extension, their motion, their duration, and perhaps a limited number of other qualities such as mass or electric charge (in some schools, most notably Descartes, these are treated as extra dimensions; in others they are treated as additional qualities of the particles in addition to extension).
(3) That matter is fundamentally indestructible and cannot be created or destroyed (except possibly by God intervening in the universe).
(4) That the only motion of significance is locomotion, the movement from one place to another.
(5) That the motions of matter can be explained according to fixed laws which operate universally and are external to both matter and independent of God (except, perhaps, as their author) which describe how motion is increased and decreased, with the same laws applying to all matter. These laws cannot be broken (except possibly by God).
(6) That these laws are deterministic, and so if one knows the positions and motions of all matter at one instant in time, then one could in principle compute the positions and motions of all matter at all earlier or later times.
(7) That space and time are independent of matter, and provide a background against which extension and motion can be measured.

There was one final postulate, which was accepted by most mechanists, but not all,

(8) The universe exists and evolves independently of God; at best God was needed to start the system off and possibly to occasionally interfere by breaking the laws of physics to create miracles.

The final postulate I believe follows from the previous postulates, but these are not necessary to reach it as a conclusion. There are certain additional perspectives on how individual elements of matter interact with each other. Some schools state that motion can only be changed by contact or collision; others allow for action at a distance (with Newton's theory of gravity as the most obvious example). There is an extension to this world-view, which states that the mechanical universe is all there is, with some mechanists opposed to this extension (Descartes being the most obvious example) and others supporting it (such as Hobbes).

Mechanism also has a strong affinity with nominalism, though not all mechanists are nominalists and not all nominalists are mechanists. If everything is reducible to fundamental parts, and covered by universal laws, and the laws acting on these parts are external nature is blind to which 'type' of matter it is acting on (as is the case with Newton's laws), then it is easy to suggest that there is no need to introduce those 'types' into our description of nature; they are superfluous. Similarities between different objects are just related to an accidental similarity in arrangements of parts. These 'types' then either only exist in our

mental classification, or can be discarded altogether, depending on which school of thought the one thinking along these lines belongs to.

Mechanism also does not need to deny the idea of purpose in nature; there have been mechanists who have added concepts of purpose to it (and some, such as – although this interpretation is controversial – Kant who have said that this addition is necessary), while others who have denied all need for purpose in discussions of nature. This second camp has a firm majority today. All schools of mechanism are opposed to the Aristotelian concepts of formal and final cause. The idea that compound objects can be reduced to the sum of their parts eliminates or at least reduces to secondary importance the role of internal structure in determining the properties of the object, and since structure is part of form this eliminates the need for formal causality. Since motion is determined through external differential equations, there are no final causes: possible effects do not arise from the nature of the being, but only from the determined actions of the external laws. The denial of formal and final causality is central to modern philosophy, and the most important feature that distinguishes it from classical philosophy.

The justification for these postulates was partly taken from theology, and partly from the successes of the early mathematical and experimental physicists, with several of the assumptions taken from the methodological premises made by those early physicists. The idea that the world was deterministic, for example, was originally taken as certain from God's foreknowledge of the future[8]. If, as classical religion demands, God knows what will happen in the future, and if we suppose that God experiences the unfolding of time as we do, then the only way that God could make prophecies is if it is possible in principle to compute the future state of the universe from a sufficient understanding of the present. This requires determinism (indeed, it is the definition of determinism). An alternative way of looking at it is that God's foreknowledge of the future leaves the future fixed; however this is a weaker condition, and does not imply that it is possible to compute the future if one knows enough about the present. There could still be random, and hence non-computable, influences on the way the world evolves, but despite this it still will evolve in one particular way. However, this perspective requires a very different concept of God, and that He does not experience time as we do as moments moving in succession.

An alternative justification for determinism does not require theological assumptions, but follows directly from the other premises of mechanism. The argument is based on two premises and one definition. Firstly that the universe is rational in the sense that there is a sufficient cause for every event. The second premise is that nature, containing the laws of physics and physical objects, is complete, in that nothing supernatural can interfere with the natural world in the ordinary course of events. These are both foundational premises of the mechanical philosophy. The definition of determinism is that physics is not indeterminate. To be indeterminate means that it is impossible in principle to predict future events to an arbitrary precision given a complete knowledge of the causes.

The universe can be indeterminate either if it is impossible in principle to have full knowledge of all causes, or if effects can occur without causes. The assumption of the completeness of physical law means that a complete knowledge of all causes follows from knowledge of the physical law (which is certainly possible in a mechanistic world-view), and of the initial conditions (which may be practically impossible, but in a mechanistic world

[8]This idea was raised by numerous Protestant theologians and philosophers (though largely rejected by the Roman Catholics); for example Hobbes wrote *To which I could add, if I thought it good logic, the inconvenience of denying necessity, as that it destroys both the decrees and the prescience of God Almighty; for whatsoever God has purposed to bring to pass by man, as an instrument, or foresees shall come to pass; a man, if he have liberty, such as his Lordship affirms, from necessitation, might frustrate, and make not to come to pass, and God should either not foreknow it, and not decree it, or he should foreknow such things shall be, as shall never be, and decree that which shall never come to pass.* (Thomas Hobbes, *Of Liberty and necessity* (1646), 36). I will later return to this point; in practice this is both bad physics and bad theology.

view is possible in principle[9]). Obviously our knowledge here is imprecise, and that will affect the precision of the predictions. But we can improve the precision of our knowledge to an arbitrary degree, and this will improve the precision of the predictions. The nature of the universe cannot depend on the precision of our knowledge, since the nature of the universe is fixed while our precision depends only on how much effort we put in. Thus reality cannot be indeterminate solely because of the vagueness or imprecision of our knowledge; if that is the only source of uncertainty, and the other assumptions hold, then the universe would be deterministic. The assumption that the universe is rational underlies all science, since without it there would be no regularities of any sort in nature to observe, and it would be impossible to make any predictions of any sort. The very success of science strongly implies that this assumption holds. If the universe were irrational, how can we trust that things like the conservation laws, such as of energy and momentum, would be satisfied? Thus it can be argued that a mechanistic understanding of the laws of physics implies that physics would be deterministic.

The idea that the universe is regular is also firmly based on three aspects of Christian theology; firstly the idea that God himself is constant,

> For I the Lord do not change; therefore you, O children of Jacob, are not consumed[10].
>
> Every good gift and every perfect gift is from above, coming down from the Father of lights with whom there is no variation or shadow due to change[11].

The second aspect is that God upholds the universe directly,

> He is the radiance of the glory of God and the exact imprint of his nature, and he upholds the universe by the word of his power[12].
>
> For by him all things were created, in heaven and on earth, visible and invisible, whether thrones or dominions or rulers or authorities all things were created through him and for him. And he is before all things, and in him all things hold together[13].
>
> You cause the grass to grow for the livestock and plants for man to cultivate, that he may bring forth food from the earth[14].

Putting these two together leads to the idea that that if the universe proceeds according to God's direction, and God's nature and thus also His direction is constant, then the manner in which the universe proceeds is unchanging. The third aspect is that God created man in His own image and likeness, with the ability to come to accurate (if not necessarily complete) knowledge of both God and the creation.

---

[9]This is even true in quantum mechanics. In quantum mechanics, we cannot have complete and simultaneous knowledge of all observable properties. But we can have complete knowledge of the wavefunction in some representation. It is knowledge of the wavefunction, not the properties, which is more fundamental, and which we use to make predictions, and thus it is this type of knowledge which is referred to in the argument for determinism from the premises of the mechanical philosophy.

[10]Malachi 3:6

[11]James 1:17. There are also additional passages which testify to the unchanging nature of God, such as Numbers 23:9, Hebrews 13:8 and Hebrews 6:17-18. Against this are passages such as Jonah 3:10 and Exodus 32:14 which speak of God relenting or changing his mind. The standard interpretation is that these latter passages are figures of speech where the language is used metaphorically. If there is to be consistency, then one of these two sets of references must be a metaphor, and the latter passages make sense as an analogy (it is perfectly reasonable to interpret them in that way), while the former passages don't.

[12]Hebrews 1:3

[13]Colossians 1:16-17

[14]Psalm 104:14. This whole Psalm is a reflection of two doctrines: that God is the underlying cause for these things, and that He did so for the purpose of providing for man and for other creatures. Atheists would perhaps retort that we now know and understand how physical principles lead to the growth of the grass. However, for theists, the two statements 'Physics did it' and 'God did it' are just two different ways of expressing the same thing, since physics is a description of how God acts to uphold the universe.

The regularity of the universe is directly confirmed in specific instances by a few passages, such as,

> Thus says the Lord, who gives the sun for light by day and the fixed order of the moon and the stars for light by night, who stirs up the sea so that its waves roar – the Lord of hosts is his name: "If this fixed order departs from before me, declares the Lord, then shall the offspring of Israel cease from being a nation before me forever."[15]

This idea of the universe being regular; that God's will is subject to His rational intellect and nature rather than God's intellect and nature being subject to His will, is a distinctive feature of both Christianity (and the Judaism from which the Christians inherited the idea) and Greek philosophy, but few other systems of ancient knowledge. For example, the traditional Persian, Islamic, Indian and Chinese religions specifically deny it; Islam in its early years experienced a conflict between philosophy and its religious nature, and after a period in which it was open to philosophy and some progress was made, gradually the religious element was victorious and Islam became intellectually moribund. However, European culture was a combination of Greek philosophy and Christianity, and thus a fertile culture for the development of science.

The third main root of mechanism was the idea that the physical world could be understood at least partially in mathematical terms. There are hints of this in certain schools of Greek philosophy, but the full discovery of this occurred in Medieval Latin Europe, one of the crucial innovations of this period and place, before being accepted and worked out wholesale by the great mathematicians of post-renaissance Europe. The fourth component, atomism, was originally due to the ancient Greeks, though again it was revived and the details were worked out by medieval Christians. Neither of these two are implied by Christian scripture, though neither is Christian scripture contrary to them, so there was no insurmountable religious resistance against these ideas.

It should also be obvious from my discussion that mechanists are not uniform in their belief; there are many different perspectives within the same general system. Many would have disagreed, wholly or in part, with a few of my eight postulates. All these thinkers would agree with most of what I have defined mechanism as, and many with all of it, but the field of mechanism is a little more varied than I have implied in the discussion so far. Indeed, the example I will discuss below is an example of that variation. My introduction here should thus only be thought of as a rough description, with not every element of it applicable to all thinkers. Each writer would have to be treated on their own terms. However, every mechanist would accept (perhaps in modified form) most of my eight postulates; every non-mechanist would reject most (though not necessarily all) of them.

The usual analogy is that of a mechanical device, such as a pocket watch, which is composed of various discrete parts joined together (perhaps by a watchmaker); these parts being arranged in a way that is not determined or suggested by their own individual nature, but which are in motion and those motions are governed by predictable and regular laws which would continue indefinitely (baring intervention from outside, which for the universe as a whole isn't an issue).

The modern mechanical philosophy might differ from this earlier form in several respects. However, in the light of Newtonian science, all the postulates I have identified are quite reasonable, and all mechanists from the time of Galileo until the late nineteenth century would have held to them (or most of them). Maxwell's discovery of the laws of electromagnetism altered how we think about the interactions of matter. In addition to the discrete particles, he (building on the work of others such as Faraday) established there was another sort of object, extended fields. What fields fundamentally are is less clear: Aristotelians would classify them as a different type of matter; I think that most mechanists

[15]Jeremiah 31:35-36; the implication being that the offspring of Israel shall not cease from being a nation. See also Jeremiah 33:20-21

would restrict matter to just the particles and say that fields are something different from matter. Nonetheless, the nineteenth century physicists showed that they had some real existence. Instead of particles interacting with particles through instantaneous collisions, we now have a picture of particles continually interacting with fields, fields with particles and fields are also influenced by other fields. Particles don't directly interact with other particles, but only indirectly, mediated through fields. This picture obviously changes the first premise I listed of mechanism; however the change is not significant enough to radically alter the philosophy, nor the conclusions drawn from it. Up to the emergence of quantum physics, then, this modified form of mechanism seemed to be fully consistent with and supportive of the best theories of physics, and most people consequently assumed, without perhaps giving it enough thought, that any reasonable philosophy would have to incorporate mechanism in some form or another.

One postulate that many mechanists agree on (though more so today than when the mechanical philosophy was first developed) is that the universe and the laws which govern it exist independently from God: at best God starts the system going, and interferes with it from time to time. However, it is easy to see how God can be eliminated from this picture (if, for example, the universe is determined to be eternal, and the evidence for miracles is denied), leading to either atheism or agnosticism. Both atheism and agnosticism obviously rely on this postulate. Generally speaking, mechanists divide the workings of the universe into the regime ruled by natural law, which proceeds independently of God, and the regime not ruled by natural law, where irregularities might occur for which we need to find some alternative explanation (such as God). Atheists argue that, as science has increased and our understanding has grown, the number of physical processes which could be placed into the region where natural law does not apply shrinks, and the most reasonable thing to say is that there is nothing in that regime; thus everything is governed by natural law, and there is no need to resort to God to explain anything. The main alternatives to mechanism (at least in Western culture) are theism, the Aristotelian theory of internal formal and final causes, and pantheism. Pantheism denies that the material universe is independent of God, but rather (in some forms) an expression or part of God; while theism agrees that the material universe is independent of God but denies that the 'laws' (it may not use that image, but has different images which take its place) operate independently of God, but rather are a description of how God sustains and upholds the universe. Theists and pantheists will also dispute many of the other postulates of mechanism.

Any Aristotelian would disagree with the first postulate of mechanism listed above, believing that substance is not reducible to the sum of its parts, but rather the parts are absorbed into the new substance and only exist virtually within it; new properties might arise from the union of these parts. They would disagree with the second postulate, saying that there is more to matter than just extension and motion, and all of the substance's nature (including its extension and motion, but not limited to these) is described by the form. They might or might not disagree with the third postulate, but will certainly see no need to suppose it, and most (I think) would disagree, seeing how one form (such as wood) can be converted into another (such as fire). They would certainly disagree with the fourth postulate, saying that change of quality within a form need not be described by locomotion. They would disagree with the fifth postulate, saying that the regularities of nature are best described by final causes internal to each substance, with different final causes pertaining to different types of substance. They might or might not disagree with the sixth postulate: Aristotelian causality need not be deterministic (since it relates substances rather than events), but is not inconsistent with it. Traditional Aristotelianism would certainly disagree with the seventh postulate (and Aristotle himself certainly would have), though some modern Aristotelians accept a geometrical conception of space and time in preference to the older view (I should here mention Leibniz, a contemporary of Newton and whose works in mathematics and physics almost reached Newton's stature, who engaged in

an early attempt to show the consistency between Aristotelian philosophy and the emerging natural science); and the last postulate would also be a matter of disagreement between Aristotelians, with theists in particular disagreeing with it. As can be seen, mechanism is wholly inconsistent with philosophies inspired by Aristotle's system; most mechanists would cite that as an advantage of their system.

Like most philosophies, mechanism has its origins with the ancient Greeks. However, it was largely suppressed by the Aristotelian causal explanations, which seemed for most of the intervening ages to be in better agreement with observation. It was revived in the high middle ages.

The idea of a mechanistic universe was taken up by the early mathematical physicists. Bradwardine had already compared the universe to a clock; Buridan explained:

> In the celestial motions, there is no opposing resistance. Therefore, when God, at the creation, moved each sphere of the heavens with just the velocity he wished, he then ceased to move them himself and since then these motions have lasted forever due to the impetus impressed on the spheres[16].

The high middle ages in Europe, the period leading up to the black death, were a time of considerable technological advancement: not as rapid as the period starting from the agricultural and industrial revolutions of the eighteenth centuries, but perhaps as much as any other period of human civilisation. There were leaps forward in agricultural technology, optical technology, weaponry, architecture, printing, textiles, and clocks. These new machines, with their intricate parts working together in harmony, together with the ongoing developments in mathematical and experimental physics, revitalised the idea of using a mechanical device as an analogy for the inner workings of the universe, contrary to Aristotle's inspiration of a living organism. This idea was accepted, to a greater or lesser extent, by most of the early pioneers of modern science.

One influential early modern proponent of mechanism was Hobbes, who was among the first to clearly formulate the principle of inertia. Hobbes saw the problems of Descartes' division between a mechanical world of extension and a spiritual world of mind, and was possibly the first person to consider that everything might be explained just in terms of bodies in motion, including our minds and thoughts. This left open the account of secondary qualities, those sensations which don't seem to be mechanical in origin, and these Hobbes declared to be either illusions or really reducible to mechanical properties through means not yet known. Hobbes was consequently a strong nominalist, believing that the common division of the world into various classes of beings was just an illusion. For Hobbes, causality can always be reduced to particular motions, and all valid explanations can be explained in terms of a reduction to elementary parts whose temporal relations can be described by a variation of efficient causality only.

As an example of an early discussion of the mechanical philosophy (I can only give one example; my intention is not to give a history of early modern philosophy, but only to outline some of the trends in the transition between medieval and modern thought), I will select Robert Boyle. By Boyle's time, mechanism was already well established by people such as Descartes (though Descartes was a dualist who disagreed that everything was mechanistic), Hobbes (who stated, contrary to Descartes, that everything was governed by mechanism), Galileo (who was less sure) and Gassendi. I will look at Boyle's *Of the Excellency and Grounds Of the Corpuscular or Mechanical Philosophy*, which was attached as an appendix to his *The excellency of theology compared with natural philosophy*. Boyle is best known as the founder of modern chemistry, and also for his work in physics on gases; his life overlapped with Newton's. *Corpuscles* mean the elemental particles of matter, similar to the concept of atoms; however, while the atomic hypothesis (as Boyle understood it) held that the atoms moved by chance while suspended in a vacuum, and it was the movements

[16]Buridan, quoted in (Hannam, 2010).

of the vacuum which directed the movements of the world, the theory of corpuscles denied that such motion was by chance, but rather by fixed laws.

I'll start with his definition of the mechanical philosophy,

> But when I speak of the Corpuscular or Mechanical Philosophy, I am far from meaning ... with some modern philosophers, that, supposing God to have put into the whole mass of matter such an invariable quantity of motion, he needed do no more to make the world, the material parts being able by their own unguided motions, to cast themselves into such a system (as we call by that name); but I plead only for such a philosophy, as reaches but to things purely corporeal, and distinguishing between the first original of things, and the subsequent course of nature, teaches, concerning the former, not only that God gave motion to matter, but that in the beginning He so guided the various motions of the parts of it, as to contrive them into the world he designed they should compose, (furnished with the seminal principles and structures or models of living creatures), and established those rules of motion, and that order amongst things corporeal, which we are wont to call the Laws of Nature. And having told this as to the former, it may be allowed as to the latter to teach, that the universe being once framed by God, and the Laws of Motion being settled and all upheld by His incessant concourse and general Providence; the phenomena of the world thus constituted, are physically produced by the mechanical affections of the parts of matter, and what they operate upon one another according to mechanical laws (Boyle, 1674).

To summarise Boyle's definition of the mechanical philosophy,

- God created matter and gave its its initial motion.
- God subsequently ordered that matter and motion in order to create the world we see and which he desired.
- Subsequent to that, God established various rules of motion which matter is forced to follow.
- All the phenomena of the world are explained by matter moving according to these laws.
- These laws are upheld and sustained by God.

This clearly differs from my eight points above, but contains numerous similarities. My first four and the sixth and seventh postulates are not explicitly stated by Boyle here, but can be assumed from the rest of his philosophy and underlie this statement. The fifth and eighth points Boyle explicitly distances himself from (though he mentions 'modern philosophers' who accept them), but only in his sense that the rules of motion are not apart from God but instead depend on his continual upholding of nature. That, I think, is the only major difference between the 'typical' mechanical philosophy I outlined above, and the mechanical philosophy of Boyle. There are also some minor differences. He states that God initially acted contrary to these laws to order the universe. He also states that this philosophy only holds to things corporeal (what I have in this work called material), leaving open the possibility of the incorporeal not governed by mechanism. One interesting thing is the similarity with the theist position outlined at the start of this chapter, with Boyle requiring God to continually uphold the laws of nature. However, this differs from the theist position in that Boyle adds in an intermediary. For the theist, God acts directly on substances, and the 'laws of nature' (or 'laws of natures') provide a description of how He does so in normal circumstances. For Boyle (at least, as I read him), God acts to sustain the laws of nature, which have their own existence independent of God, and it is these laws which influence matter.

Note also that Boyle is among those who believe that motion can be changed only through direct contact: the action at a distance of magnetism or gravity needs to be explained in terms of intermediaries. For this reason, Boyle proposed the idea of the aether, an invisible fluid permeating all space. The particles of this aether could form various flows or whirlpools which affect the motions of the various visible bodies moving around in them. This is a natural requirement of the idea, adopted by the mechanism favoured until Newton (indeed, which seemed to be an inevitable consequence of the mechanistic ideas), that motion could only be transferred through physical contact, and was needed to explain phenomena such as gravity and magnetism. Boyle considered that there might be two types of aetheral matter, one to explain gravity and the other magnetism.

But what interests me most is Boyle's reasons for preferring his version of mechanism over other systems. He gives these as

(1) The mechanical philosophy is clear in both its principles and how these lead to explanation of phenomena. Both Aristotelians and alchemists are very unclear in principles, and thus have many disputes, and are even more unclear when it comes to applying those principles to matter.

> And even in some of the more ingenious and subtle of the Peripatetic discourses upon their superficial and narrow theories, me thinks, the Authors have better played the part of painters than philosophers, and have only had the skill, like drawers of landscapes, to make men fancy, they see castles and towns, and other structures that appear solid and magnificent, and to reach to a large extent, when the whole piece is superficial, and made up of colours and art, and comprised within a frame perhaps scarce a yard long.
>
> ... Those very Aristotelians, that believe the celestial bodies to be moved by intelligences, have no recourse to any peculiar agency of theirs to account for eclipses. ... For no intelligent man, whether alchemist or peripatetic, flies to his peculiar principles, after he is informed, that the Moon is eclipsed by the interposition of the Earth between her and it, and the Sun by that of the Moon between him and the Earth. And when we see the image of a man cast into the air by a concave spherical looking-glass, though most men are amazed at it, and some suspect it to be no less than an effect of witchcraft, yet he that is skilled enough in optics, will, without consulting Aristotle, or Paracelsus, or flying to hypostatical principles and substantial forms, be satisfied, that the phenomenon is produced by the beams of light reflected, and thereby made convergent according to optical, and consequently mathematical laws. (Boyle, 1674)

(2) The reasons that opponents give against mechanism is not that their explanations or descriptions aren't clear, but that they only apply to a few applications.
(3) There cannot be fewer principles than those of the mechanical philosophy, namely matter and motion.
(4) Nor can we conceive any principles more primary than matter and motion, since they must have both been immediately created by God together at the first.
(5) Neither can there be any physical principles simpler than matter and motion, neither of which are compound qualities.
(6) The mechanical principle is comprehensive. When matter collides, the only possible effects are that it is broken up, into particles of determinate motion, figure, size, posture, rest, order or texture; or to drive its entire bulk. These actions are those covered by mechanical theories. Furthermore, there is a huge number of variations produced by the composition and decomposition of such matter, all contained by myriads of single invisible corpuscles. With this huge number of possible bodies and phenomena available, there can be no doubt that mechanical principles can

explain all phenomena of natural philosophy; people who say that it works well for some phenomena but not others are like those who say that because the Latin alphabet may able to describe the words of one book such as Virgil or Euclid, but that is no reason to suppose that it cannot be used for every book in every language in the library. As for those who would say that the mechanical principles are not sufficient, and we need to add to this account substantial forms, real qualities and similar non-mechanical principles and agents:

> But this is not necessary; for, both the mechanical affections of Matter are to be found, and the Laws of Motion take place, not only in the great masses, and the middle-sized lumps, but in the smallest fragments of matter; and a lesser portion of it, being as well a body as a greater, must, as necessarily as it, have its determinate Bulk and Figure: and he that looks upon sand in a good microscope, will easily perceive, that each minute grain of it has as well its own size and shape, as a rock or mountain. And when we let fall a great stone and a pebble from the top of a high building, we find not but that the latter as well as the former moves in conformity to to the laws of acccleration in heavy bodies descending. (Boyle, 1674)

Equally, those who cite the likes of substantial form still need to provide the mechanism that allows it to interact with matter. What is the general agent that produces a phenomena and by what means and in what manner does it do so?

> So that the mechanical philosopher being satisfied, that one part of matter can act upon another but by virtue of local motion, or the effects and consequences of local motion, he considers, that as, if the proposed agent be not intelligible and physical, it can never physically explain the phenomena; so, if it be intelligible and physical, it will be reducible to matter, and some or other of those only universal affections of matter, already often mentioned. (Boyle, 1674)

(7) The mechanical philosophy is so universal that it must include any alternative explanation of nature. Every success that these ideas have could also be deduced from mechanical principles, while every failure is obviously of no account.

> From the foregoing discourse it may (probably at least) result, That if, besides Rational Souls, there are any Immaterial Substances (such as the Heavenly Intelligences, and the Substantial Forms of the Aristotelians) that regularly are to be numbered among Natural Agents, their way of working being unknown to us, they can but help to constitute and effect things, but will very little help us to conceive how things are effected; so that, by what ever Principles Natural things be constituted, it is by the mechanical principles that their phenomena must be clearly explicated. (Boyle, 1674)

(8) There is an analogy that so many diverse buildings can be built out of the same bricks, so a great diversity of macroscopic objects built out of the same corpuscles.

(9) The success of the mechanical philosophy in many fields leads to the hope that it will eventually succeed in all fields.

(1) Of the principles of things corporeal, none can be more few, without being insufficient, or more primary than Matter and Motion.

(2) The natural and genuine effect of variously determined Motion in portions of Matter, is, to divide it into parts of differing sizes, and shapes, and to put them into different Motions, and the Consequences, that flow from these, in a World framed as ours is, are, as to the separate fragments, posture, order, and situation, and,

as to the Conventions of many of them, peculiar Compositions and textures.

(3) The parts of Matter endowed with these universal affections are by various associations reduced to Natural Bodies of several kinds, according to the plenty of the Matter, and the various Compositions and Decompositions of the Principles; which all suppose the common matter they diversify: And these several kinds of Bodies, by virtue of their motion, rest, and other mechanical affections, which fit them to act on, and suffer from, one another, become endowed with several kinds of qualities,... and those that act upon the peculiarly framed organs of sense, whose perceptions by the animal faculty of the soul are sensations.

(4) These principles, matter, motion, (to which rest is related) bigness, shape, posture, order, texture, being so simple, clear, and comprehensive, are applicable to all the real phenomena of nature, which seem not explicable by any other not consistent with ours. For, if recourse be had to an immaterial principle or agent, it may be such an one, as is not intelligible; and however it will not enable us to explain the phenomena, because its way of working upon things material would probably be more difficult to be Physically made out, than a Mechanical account of the phenomena. And ... we cannot conceive how it should produce changes in a Body, without the help of Mechanical Principles, especially Local Motion ....

(5) And if the agents or active principles resorted to, be not immaterial, but of a corporeal nature, they must either in effect be the same with the corporeal principles above-named; or, because of the great universality and simplicity of ours, the new ones proposed must be less general than they, and consequently capable of being subordinated or reduced to ours, ... (Boyle, 1674)

So, having seen the argument set out at length, how do we evaluate it? The reason why I cited it at length was to make an important point, which is related to an omission made by Boyle: nowhere does he show the scholastic position to be inconsistent, or its premises false, or prove it incapable of accounting for the phenomena he is interested in studying (and neither, to my knowledge, did any other renaissance or early modern scholar). He argues that the mechanical principles can, and obviously they are inconsistent with the scholastic principles, but this is not to say that the scholastic principles cannot also do as well; neither is it that nobody of the time explained these phenomena using scholastic methods as his own an argument that it is in principle impossible. For example, suppose that a consideration of efficient and final causes can be adapted to explain the phenomena of nature as well as the mechanical principles? Boyle responded by saying that if so then they could be absorbed into the mechanical picture, but why not take it the other way, and say that the mechanical view can be absorbed into the scholastic system?

Equally, it is important to note that Boyle's argument is only universal if the only problems of interest are those related to matter, motion, bigness, shape, posture, order and texture. A modern scientist would, of course, add qualities to this list (including mass and electromagnetic charge), but this does not overly affect Boyle's underlying principle, since these qualities can also be treated (or so it seems) by the mechanical philosophy. But it still lends to an issue of begging the question: Boyle accords an argument success if it can account for these and related qualities; when he regards an idea as successful it is done by comparison with this reduced standard that he derived based on its affinity with what mechanics does well. In other words, Boyle is excluding everything for which mechanics is not suited to explain before he starts. This is akin to designing the rules of the game to

ensure that the Germans always win. There is, it is claimed, far more to knowledge than physical phenomena, and Boyle does nothing to dispute this claim, beyond an appeal. To be fair to Boyle himself, he recognised this and acknowledged it: he did not claim that all qualitative explanations are useless (which the likes of Descartes and Hobbes did claim). His criticism of teleological explanations was not that they were invalid, but that the question of *how* (to be answered by mechanical principles) should be given priority over the question of *why* (which leads to a telcological explanation). This is, in my view, not too unreasonable: before answering why something is, it is vital to first understand what it is, and even more so that it is.

Thus the idea of delaying the answer of *why* until one understands the phenomena well enough is reasonable; but there is, of course, the question or whether mechanical principles can answer this question, or if there is an answer to this question. For example, one can state that acceleration is force divided by mass, and that all bodies attract each other with a force proportional to the product of their masses, and inversely proportional to the square of the distance between them, and this will (to a very good approximation) allow one to understand how the solar system evolves. But it doesn't provide a complete explanation, for that would include an answer to answer why Newton's various laws apply to matter. Of course, we can go into more advanced physics than Newton, but that doesn't answer the underlying question, but just changes it to a better posed (or more relevant to the real world) inquiry. Expanding the mathematical picture doesn't help, because the question is why does the mathematical picture work at all, and what are its limitations, and is there anything that can't be expressed mathematically? These questions cannot be answered by mathematics; we need the mathematics to tell us what the question is, but that is as far as we can go using the methods of modern physics; we need something beyond mathematics to answer it. At some point, we are going to have to go beyond physics and the mathematical explanation. Of course, we need to get the physics right first.

However, if the idea of delaying the why questions until the *how* questions are answered is reasonable, we must ask whether the early mechanists did that; since mechanism makes numerous presuppositions about the nature of physics; presuppositions which are used in the interpretation of physics but not derived from physics. This is not the fault of mechanism; it is true for any system which interacts with science: we must presuppose certain things before we interpret even our first observation, and thus those presuppositions cannot be derived from experiment but come before it. Of course, a contradiction would lead to those presuppositions being questioned; however, a lack of contradiction doesn't show them to be true. The mechanical system is no less of a metaphysical idea imposed on experiment than the scholastic system.

Boyle assumes without a strong enough argument that matter and motion are both necessary and sufficient to understand the physical world. Clearly, if this assumption fails (for example if the structure or state of matter is also important), then so would Boyle's system. Boyle is wrong to say that only the mechanical system is concerned with matter and motion, for example the Aristotelian system is concerned with these, albeit defining motion in a broader sense (and should it be found that movement from place to place is not the only type of change important to physics, then the more general Aristotelian system will clearly be able to cope with that better than the more limited mechanical philosophy). He is also wrong to say that the mechanical philosophy is built on only matter and motion, because it adds in concepts such as forces, and (later on) fields, not to mention the entire concept of a law of motion. That Aristotle added to matter and motion the ideas of form and finality as fundamental is no criticism of his system, since the concepts of matter and motion by themselves are not sufficient to constrain a philosophical system.

Boyle's particular examples of the failures of Aristotelian physics are also strange. For one thing it is Aristotle's metaphysics, as it was developed by the medieval philosophers, which is the rival to mechanism rather than his physics; Aristotle's metaphysics is consistent

with many systems of physics, not just the one which he himself proposed. Secondly, that eclipses are caused by the moon interposing itself between the earth and the sun was discovered by the classical philosophers. The existence of lenses and mirrors, and the laws of refraction that describe the behaviour of light passing through lenses, are no barrier to the Aristotelian system. One can discuss those laws without reference to anything but the equations, but once one tries to understand why those laws hold then one is forced into a discussion of the nature of light and glass, and it is here that the Aristotelian would invoke form, and the mechanist his own principles: the facts of refraction and reflection, which is all Boyle uses, is consistent with both the Aristotelian and the mechanical philosophies. The disagreement between Aristotelian metaphysicists and mechanists is not over the facts of the experimental data, but over why those regularities exist in nature.

Boyle's statement that the mechanical system is comprehensive is unproven. It asserts that the only motion of significance to physics is locomotion; if this is true then the mechanical system might well be comprehensive. But the truth of this statement can only be established as a conclusion of studies into physics, once we have the final physical theory, and not as a premise made at the start of modern physics. Equally, as I mentioned earlier, his statement that any other system would be absorbed into mechanism has things the wrong way round: the more specialised system should (if they are compatible) be absorbed into the more general one. Aristotelian metaphysics is more general than mechanism, because it can describe other forms of change than just locomotion (although if locomotion were the only type of change the Aristotelian system could cope with that). By adopting mechanism one would run into the risk that physics might require some other form of change, and thus have reality conflict with a premise; adopting a more general system avoids that risk. Of course, in practice, mechanism and the Aristotelian philosophy would not be equivalent even if they showed perfect agreement in explaining experimental observations, since they are built from incompatible premises. Thus Boyle's contentions that mechanism is complete is unproven and that the alternative system would at best reduce to mechanism is false.

But the biggest potential weakness of Boyle's argument is an assumption that underlies his whole thesis: that a physical law consistent with the mechanical principles can explain all of physics. If this assumption fails, then everything he wrote collapses. This is not something he proves; he merely takes it to be obvious. Now, in Boyle's day this assumption seemed quite reasonable, given the successes of mechanical physics, and he seemed to be vindicated by the success of Newtonian physics, which makes the same set of assumptions. But Boyle lived at the start of the scientific revolution, not at its culmination; his (and Newton's) knowledge of physics was (by today's standards) rudimentary. Thus this assumption should not be taken for granted. It needs to be established in terms of the most fundamental physical theory before it can be wielded.

So Boyle, like all the early mechanists, failed to offer any fool-proof arguments against a more Aristotelian metaphysics. There was no first-principles refutation of the Aristotelian system (including its premises) or demonstration of its inconsistency; and he failed to demonstrate that the Aristotelian philosophy could not give as a good an account of the observed phenomena as the mechanistic system. The best that he could say was that the mechanistic system was successfully describing the physical processes thus far studied to within the precision of the contemporary experiments, which is a long way from saying that it is true. Could one argue, as Boyle did, that because mechanism was successful in inspiring theories that explained all experiments that had been conducted by the time he was writing that it would also be successful in all future experiments? There might some indication, but it is not a rigorous argument. However, Boyle's version of mechanism does have the advantage that it can be falsified: by reducing physics to the study of matter and locomotion, he opens the door to his philosophy being experimentally disproved.

Another issue that was being faced by the early modern philosophers was the question of how we can know. If there is a division between the soul and the mechanical body, then

what is the connection between the two, and how does our sensual input influence us? Boyle himself was pessimistic about how much humanity could know, being confined to the dark prison of the body. This meant that we should not reject things because they transcend our intelligence; we must always be conscious that we might not be able to explain or understand everything. This is no excuse for not trying to understand as much as possible, or to enlarge our understanding. From this, Boyle concluded that it was pointless to try to construct great and general systems: better to have a little certain knowledge than a lot of speculation. In much of his work, Boyle is satisfied to merely gather facts rather than speculate. He considered those who tried to construct grand systems as worthy, but was always wary that some new experiment usually came along to rob them of their success. In this way, Boyle always placed experiment at the heart of the new science, with grand systems standing or (more usually) falling based on how well they fit experiment, and in how well their predictions line up with reality. There may be several hypothesis which explain the data: in this case, one cannot choose between them. And, of course, that means that we cannot choose between our own favoured hypothesis, and the one we don't yet know about.

Boyle was obviously devoutly religious from what we have cited so far. He believed that God was necessary firstly to set up and uphold the rules of motion, and also because the order, vastness and beauty of the universe implied an intelligence designing it. Having thus accepted God, Boyle was content to interpret Him according to the principles of orthodox Christianity, as outlined in Holy Scripture. Science and theology are in this way two parts of a larger whole which transcends them both by far; they are each only a small part of the whole system of God's contrivances. He also considered the possibility that God could have created other universes than this one, with very different creations than those we currently admire.

Equally God was seen as essential in keeping the universe in order without God, a mechanical universe without God would inevitably fall into a kind of chaos or confused state of shuffled and depraved things. God is necessary to uphold the constancy of the law and to ensure that matter obeys it; without God there is no reason for either of these to occur. A law, to Boyle in some moments, was the sign of how an intelligent agent regulates his actions. Inanimate bodies are incapable of understanding what a law is, and thus the actions of such bodies cannot incite or moderate their actions, but are produced by Real Power rather than laws. God is thus not only the first cause, but also ever-active in sustaining the universe. God's knowledge reaches at once, to all that it is possible to know; his penetrating eyes pierce the whole creation in one look. That God framed general laws does not contradict the idea of any true miracle, since the idea of miracles presupposes an ordered and settled course of nature, but allows the author of that nature to suspend, alter or contradict those laws, which He alone first established and continually upholds, as He sees fit. God can thus invalidate the axioms of natural philosophy if He so chooses. Such miracles include those of the Bible, the joining of a rational soul to a body, answers to prayer, and, of course, any irregularities we know of (Boyle cited sunspots as an example) or don't know of. However, the main argument for God is not derived from the various irregularities in the universe, but instead from the far more common regularities: the laws of nature point to a supernatural lawgiver, and one that has a profound appreciation for mathematical beauty.

6.7. Astronomy

Although we now naturally view astronomy as a subsection of physics, it was seen as a separate discipline until Newton unified earthly and heavenly dynamics. The rules that governed the planets and stars were *obviously* different than the rules that governed motion on earth. Astronomy was also different from physics in that it was seen since ancient times as being closely related to geometry. There had always been an interest in the motion of the planets, principally to provide for the needs of the Astrologers, who needed to be able to compute both the birth charts of people and predict the paths of the planets to perform their

fortune-telling. Although the church (mostly) frowned on it[17], it nonetheless had moved into Europe alongside the (more useful) examples of ancient Greek and Islamic learning[18], and had a wide following. Thus there was always a market for more accurate observations and theories of the heavens, and this proved beneficial. The main task of the astronomers was to map out the paths of the planets – why they travelled along the trajectories that they did was not questioned, as it had been answered by Aristotle: heavenly bodies were made of quintessence, which had the natural tendency towards circular motion. As we have already seen, this was expanded by later Medieval writers such as Buridan, who thought that the theory of impetus might not also be applicable to the planets. Thus, astronomy was all about mapping out paths, which meant geometry.

In the ancient world, there had been many theories of the motions of the planets, including several which put the sun at the centre of the solar system, but Ptomely's theory came out on top because it was by far the most successful in matching with the observations. Ptomely placed the earth at the centre of the universe, and stationary (i.e. not rotating) and introduced a system of epicycles to explain the motions of the planets. This system involved having the planets in a circular orbit around some (unphysical) point which was in a circular orbit around some other point, until eventually you get a circular orbit around the earth. This system was ungainly, but it seemed to match observations perfectly.

It was known that there were numerous scientific difficulties with a heliocentric system (the sun fixed at the centre of the solar system). This system required a rotating earth, which seemed to have various objections against it: why were not we propelled by the fast velocity of the earth into space? Why was not there an enormous wind as we passed through the air at an huge speed? Why, when we shot an arrow into the air, did not it fall behind us (or in front of us, depending on which way we were facing) as we moved with the earth and the arrow just engaged in vertical motion? Why did it seem to our direct experience that we were stationary, if we were really moving? These, as we have seen, were answered by the mathematicians of the middle ages, who concluded that we would not be able to distinguish between a system where we rotated and the universe was still, or the universe rotated and we were still. There was, however, one additional problem which was not resolved at that time: the stellar parallax. As the earth moves against the fixed background of the stars, they will appear at a slightly different angle from North. The early astronomers computed the distance of the earth from the sun (accurately enough), and thus the radius of the earth's orbit; they calculated the distance from the earth to the stars (wrongly) by comparing the apparent size of the stars against that of the sun; they calculated what the parallax should have been (wrongly, but only because they got the distance wrong); deduced that this number should have been measurable according to the precision of their instruments (rightly); and completely failed to measure it. The obvious conclusion was that the earth was stationary in relation to the fixed background of the stars. There were also various theological reasons raised against a moving earth. One example was Psalm 94:1

[17]The early Christian writers such as Augustine dismissed it entirely; in the Medieval period some scholars such as Aquinas tried to accommodate it a little. With regards to people, Astrology was always seen as either useless or demonic, because it contradicted the important doctrine of free will – even before one brings in objections from physics. The writers hedged their bets about whether it could be used to predict natural events such as the weather and volcanoes, which did follow regular laws. Of course, we now know that there is no relationship between them (apart from eclipses and tidal effects), but in the middle ages this was not sure, and a lot of people took astrology seriously, so the church wisely declined to comment as it was outside its area of authority. If only it had the same attitude when Galileo came along!

[18]Astrology originated in ancient Babylon and Persia; it continued in these regions after the Islamic conquest and was spread alongside the rest of Persian philosophy; however Islam opposed it even more than Christianity did and it was gradually suppressed in the Middle East as Islam started to dominate Middle Eastern intellectual thought.

> The Lord reigns; he is robed in majesty; the Lord is robed; he has put on strength as his belt. Yes, the world is established; it shall never be moved.

This, however, was not a serious concern. As is clear from the rest of this verse, this Psalm is full of allegorical images. It is not too much of a stretch to say that that final line should also be interpreted allegorically. 'It shall never be moved' might refer to our appearance of the world, but more likely to God's control over the world.

I should mention that the ancient view of the universe did not do so to place man at the centre of the universe, with everything else revolving around us. Rather it placed man in the worst place in the universe. The heavens were perfect and unchanging; only the earth was corrupt and liable to degeneration. Thus if we were to argue in terms of mankind's view of his place in the universe, the heliocentric view represented a promotion for humanity rather than a degradation. But of course, ultimately the debate between Ptolemy's, Copernicus', Tycho's, Kepler's and the the other models of the solar system was not about mankind's place in the universe, but merely a question of how the solar system worked and which view fit best with observation.

Against this background, Copernicus made his appearance. He had studied the ancient Greeks, and knew that there were many rival theories to Ptomely's. He had studied the scholastic mathematical natural philosophers, and knew that the most obvious objections to the idea of a rotating earth and stationary sun failed. So he naturally asked himself if Ptomely's theory was the only possible one, or indeed the best. His main motivation was a desire for mathematical simplicity: he believed that the system of epicycles was ungainly, and God, being the ultimate geometer, surely could have come up with a better way of arranging the heavens. He observed that many of Ptomely's epicycles could be removed by having the earth rotate and orbit the sun. However, the best he could do was to reduce the number of epicycles from about eighty to a little over thirty. The second reason that Copernicus offered for his system was that it was fitting that the sun, as the source of light and heat, should also be the centre of the universe; in the same way that God the Father, as the source of life and all, was also at the centre of the universe in another sense. However, he had nothing to offer against the objections to his system, both scientific and theological. With regards to the parallax, he could only suppose that the stars were much further away than had been measured based on their apparent size. Not surprisingly, his thesis was widely discussed, but rejected by almost everyone; it was regarded as a joke because it clearly contradicted observation. The only exception to this were a few mathematicians, who were swayed by Copernicus' appeal to (relative) mathematical simplicity.

The next important figure was Tycho Brahe. Like most people, he rejected Copernicus' system, as it was incompatible with observation, but unlike most people he was intrigued by its comparative mathematical simplicity. He was a Danish nobleman who made his living through astrology and sponsorship from the King in Denmark and later the Emperor at Prague as well as relying on his ancestral wealth, although he also practised alchemy and is best known today for his detailed astronomical observations. He first shot to fame by refuting Aristotle's idea that the heavens were perfectly unchanging through careful observations of a supernova. He proved that this was not just a comet through his careful measurements. He produced the most comprehensive and detailed charts of the planets yet assembled; all the more impressive because he was just a little bit too early for the telescope. He sought to merge Copernicus' ideas with Ptolemy's system, and came up with an alternative model of the solar system. Like Ptolemy, he placed the earth at the centre of the universe and stationary. For him, the issue of the parallax was crucial; he computed that if Copernicus was right and the stars were too distant for a parallax to be measured, they would all have to be larger in radius than the orbit of the earth; if further away than this, then even larger still. This seemed to him to be absurd: why should the sun be so much smaller than all the other stars? However, his careful measurements were enough to disprove Ptolemy: Ptolemy

predicted that Mars was always further from the earth than the Sun is, but Tycho's detailed measurements showed that this was incorrect. However, Copernicus' system got this right.

So Tycho, wanting to combine the Copernican system with a stationary earth, came up with an alternative. He had the sun and moon rotate around the earth, while the planets in his system rotated around the sun. Once again, he did not remove all the epicycles, just the largest of them: they were still needed to explain the deviations of the planets from a circular orbit. Tycho's system was basically Copernicus' projected into a frame of reference centred on the earth. It had all the advantages of Copernicus' system: it was mathematically simpler than Ptolemy's, explained his and Galileo's (later) measurements, but without the disadvantages. Indeed, if we make a small adjustment to Tycho's system and allow for elliptical orbits, then it is as far as the dynamics of the solar system is concerned, mathematically equivalent to the Kepler/Newton system, and with a few more tweaks can be made identical to Einstein's. The only way (as far as I am aware) that the idea of a Tycho-styled system with a stationary Earth can be experimentally distinguished (at least in Newton's physics; I think that Einstein's should be able to cope with this, though haven't done the calculation myself) from one with a stationary sun is from the phenomena of stellar aberration: because of the finite speed of light, the direction in which it appears to be omitted from a light source when the observer is in motion is slightly different from what one might naively expect. Mathematically, of course, the sun-centred solar system is much simpler and more elegant.

In terms of the history of ideas, Tycho's system was important as an intermediary step between Ptolemy and Kepler. Indeed, even if one adopted Tycho's system of the stationary earth, computations were easier if one mathematically switched to the corresponding heliocentric model, did the hard work, and then switched back to the reference frame of the observer. It allowed people to gain computational experience in heliocentric systems without having to commit to them philosophically.

Johannes Kepler was an outstanding mathematician (who even apart from his astronomical work made several important advances, such as work on the theory of infinitesimals, which was one of the precursors to calculus), who had an outstanding slice of good fortune: he was an apprentice to Tycho in the last years of Tycho's life, and was the one selected to complete and publish Tycho's astronomical tables. Kepler was also a convinced Copernican from his youth. Once again, he was attracted to the mathematical elegance and simplicity of Copernicus' system, and liked the idea of the sun at the centre of the universe with the planets depending on it and bound to it because of a third force because of an analogy between the Christian Trinity of Father (Sun), Son (planets) and Holy Spirit (the bond between the two). He expected God to leave marks of his own nature in the cosmos, and this reflection of the nature of God in the solar system dynamics was very appealing to him. Indeed, he was convinced that there were numerous mathematical regularities in the universe waiting to be discovered: his diligent studies led to his laws of planetary motion; and although some of his findings were mere coincidences, others were important in inspiring Newton.

He first came to prominence through a work that polished many of Copernicus' rough edges; then he moved to Prague to work with Tycho Brahe. This provided access to Tycho's detailed tables, and eventually (after numerous failed trails) Kepler was able to deduce from them that Mars moved in an elliptical path around the Sun with the sun as one of the focuses of the ellipse. Further studies of the orbit of Mars led him to generalise to his laws of planetary motion: that all planets move in ellipses with the sun as a focus; that a line segment joining a planet and the Sun sweeps out equal areas during equal intervals of time; and that the square of the orbital period of a planet is proportional to the cube of the semi-major axis of its orbit. This removed epicycles entirely for the first time, and matched Tycho's observations better than any other system. Mathematically it was the simplest and most elegant solution to the problem of planetary motion yet devised. Following the discovery of the earth's magnetic field, Kepler suggested that the sun's magnetic field drove

the orbits of the planets. Although it was wrong, this idea suggested to merge astronomy and earthly physics: another dangerous threat to the Aristotelian system.

Kepler was a devout Protestant, which had several advantages and disadvantages for him. It caused him some personal difficulties when employers and patrons objected to his beliefs, and also meant that, as a heretic, he was almost completely neglected in the debates in Italy. Even though he and Galileo kept a close correspondence, neither Galileo nor his opponents referred to him: after Galileo's and Tycho's observations had ruled out Ptolemy, their debate was between Copernicus' system and Tycho's. Kepler did, however, use his studies on optics to improve on Galileo's telescope. However, he was also free to ignore the proclamations of Rome, and his opposition to these increased his appeal in Protestant countries. His work on Astronomy was not widely accepted at the time, though a number of astronomers read his writings and tested his system, which was generally found to work well, and it was especially well-received in England. Kepler's explanation of the planetary motions was rejected, and others, such as Hooke, began to search for an alternative, such as an attractive force between the sun and the earth.

The main philosophical advance of the new astronomy was that it put a serious dent into Aristotelian physics, and thus Aristotle's reputation as a whole. There was little doubt that Aristotle's treatise on the heavens was in error: the universe beyond the earth's atmosphere was not unchanging; the heavenly bodies did not have a natural tendency towards circular motion because (if nothing else) they did not move in circles. Instead, their motion was well described by a mechanical law. Equally, it gave importance to the idea of searching for mathematical simplicity and elegance as a motivating factor in searching for astronomical laws. There were also suggestions that the same laws governed the heavens and the earth, which favoured the idea that phenomena on the earth should also be described by mathematically elegant mechanical laws. If Aristotle/Ptolemy were wrong about the heavens, could they also not be wrong about the earth?

As far as the Roman Catholic church was concerned, this was all bad news. The church had for several centuries fully absorbed Aristotle's philosophy into its own teachings and doctrine. This synthesis was already under attack by the Protestants, some of whom challenged various aspects of Aristotle's metaphysics as they were applied by the church, particularly with regards to transubstantiation. Now Aristotle's physics was also under attack, and while this was less significant to church doctrine than the metaphysics, it was still undermining the authority of the man who the Catholic synthesis had built upon. The church should have kept a neutral position; perhaps in other times it would have done. But not in this age. The late medieval and renaissance Catholic church was corrupt, arrogant, badly fallen from its original purity and already in a bad mood on account of the criticisms of the early protestants from Wycliffe and Huss through to Luther. Most ordinary parish priests and monks were good, decent and pious men who genuinely practised what their religion preached, but at the centre of the institution was a mess of greed, corruption, stupendous wealth, superstition, pride and immorality. It had enormous temporal power still, but that power was based solely on its intellectual prowess and claim to be the sole divinely appointed interpreter between God and man. It was not going to let go of it without a fight.

6.8. Descartes

Descartes is primarily seen today as a philosopher and the founder of modern philosophy; he saw himself as first and foremost a mathematician. A contemporary of Kepler and Galileo, he did not maintain as close correspondence with them as they did with each other, but he was aware of and partly inspired by their work (though he also rejected some of Galileo's views), his ideas had certain similarities with theirs, and the three of them together were the basis of what arose in the next generation.

Descartes fell in love with mathematics as a young man, and mastered it at an early age. His first major achievement was to invent the field of analytical geometry, which posits a one to one correspondence between geometry and space. The basic structure of this is familiar more modern mathematicians; to me it is difficult of thinking of doing geometry without it. The basic idea (in two Euclidean dimensions) is to select two non-parallel lines (they need not even be straight, although straight lines are most convenient and will always give a unique answer), and mark off equal distances along the line. Each point in the geometric space can be found by drawing a perpendicular from the coordinate axes to that point, and marking the distance where that perpendicular crosses the coordinate axes. It is easiest if the two coordinate axis are perpendicular to each other, and these are known as Cartesian coordinates after Descartes. One can extend this system to more than two dimensions easily, by just adding an extra coordinate axis for each direction. In this way, one can map every point in space and time to a set of numbers, with as many numbers as there are dimensions. This set of numbers is known as a vector. Every shape can be mapped to a sequence of these vectors; for example a triangle can be described by the three vectors that give the coordinates for its corners. Also note that there is no difficulty in extending this to an arbitrary positive integer number of dimensions[19]. We are familiar with zero dimensions (a point), one dimension (a line), two dimensions (a surface), and three dimensions (a solid); the fourth dimension (time or succession) can be added to this as well. Beyond that, we cannot imagine. The reason we cannot imagine it is because our imagination builds on our spatial awareness, and our spatial awareness is limited to the three space and one time dimensions of the world that we live in. But once we have converted geometry into number, there is no difficultly in extending geometrical concepts to five, ten, or even twenty six dimensions. We just need to write down twenty six numbers for each point, and can apply (almost) exactly the same mathematics that we do in three dimensions. The key thing to realise is that imagination is a hindrance rather than a help to our mathematical studies; we are best in leaving it behind and just use the abstract numbers and symbols. At the end of the calculation we then convert back from the abstract notation to a more physical picture.

Physical operations can be described by operators acting on these vectors. Consider, for example, rotations. We know how to rotate a physical object: we just take it in our hands, and turn it round. In the numerical representation of space, what we do is perform an operation on the vectors that described that object and for each vector we move them to a different point in space. In other words, to represent a rotation, we need a procedure which takes in one set of vectors and spews out a different set of vectors. We call this procedure an operator. So if $\vec{\mathbf{v}}_{in}$ represents the set of numbers describing our input vector[20], and $\vec{\mathbf{v}}_{out}$ the set of numbers representing our output vector. The operator is represented by a symbol $\hat{O}$, and we can write $\vec{\mathbf{v}}_{out} = \hat{O}[\vec{\mathbf{v}}_{in}]$. For certain procedures, and rotation is one of them, $\hat{O}$ can be represented by a matrix, which is described by a set of $d \times d$ numbers which lead to each component of $\vec{\mathbf{v}}_{out}$ depending only on the sum of the components of $\vec{\mathbf{v}}_{in}$ multiplied by the coefficients of the matrix according to a fixed rule. Most geometrical transformations can be represented by a matrix multiplying a vector combined with the addition of vectors; sometimes we need to be more sophisticated than this and go beyond matrices, but similar principles apply. In this way, every geometrical operation has a one to one mapping to an algebraic operation, which is good because algebra is easier than pre-Cartesian geometry. This, of course, allows us to apply the full power of mathematics to many physical problems.

---

[19]Extending Cartesian geometry to a non-integer number of dimensions is harder, but also possible. Indeed, this is one of the key methods that every quantum field theorist needs to have available to them.

[20]Note that we are extending algebra here: in algebra usually one symbol is used to represent one unknown number; here we are using one symbol to represent d numbers in d-dimensional space – so two numbers in two dimensional space, three numbers in three dimensional space, 26 numbers in 26 dimensional space, and so on. The bold face and arrow indicate that the quantity is a vector.

Of course, the Cartesian system is not uniquely defined. Which coordinate axes do we draw? Where do we place the zero of our coordinate system? How do we define the scale used to measure lengths along these axes (for example, do we make measurements of distance in meters or inches)? These choices are entirely arbitrary. A crucial point – which was perhaps not fully appreciated until Einstein, although Galileo, Descartes and Leibniz certainly had a hint of it – is that the real physical world cannot depend on these arbitrary choices which we use to construct a mathematical representation of it. The procedure is that we take a physical system, map certain aspects of that physical system to a geometrical space, choose a coordinate system in some way and use it to map from the geometrical system to an algebraic and numerical system, solve for the various laws of physics acting on the algebraic system, and then map those results back from the algebraic system to the geometry, and then from the geometry to the physical world, where we can compare the results against observation. The physical object we are studying, of course, proceeds directly from initial to final state without needing the mathematical abstraction. Neither the starting point or ending point of the calculation depend on the abstraction (whether algebraic or geometrical); only the intermediate steps which are used to convert from a physical system to an algebraic system. The final answer we get is not and cannot be affected by which coordinate system we choose. Only a few possible mathematical descriptions of physical law lead to predictions which are independent of this arbitrary choice, which means that the true physical theory must be one of these - and we can say this even before we turn to experiment (if we assume that some aspects of physics can be treated geometrically).

Note that this procedure requires that the mapping from the physical world to the algebraic system is reversible. This means that the physical system must contain as least as much information as the algebraic system, and it requires that every aspect of the initial state and final result of the algebraic computation should correspond to something in physical reality: the mathematical calculation cannot introduce elements which cannot be mapped back to physical reality if the method is to remain valid. While intermediate steps in the algebraic calculation might in principle not correspond to physical objects, in practice this is hard to achieve without introducing non-physicality into the final result. Therefore (given the success of this method in understanding those aspects of physical reality to which it is applicable) it is very likely that every aspect of the algebraic computation corresponds to something that exists either in the physical world or our knowledge of the physical world as mediated through our senses and taking into account the intrinsic uncertainty of our knowledge. It might be a one-to-one mapping, with every aspect of the physical world represented by some mathematical symbol. However, it might also be a selective mapping, with only some aspects of the physical world representable algebraically. The scientific method by itself cannot distinguish which of these cases is correct; it cannot say that every aspect of reality is accessible scientifically; nor can it say that it isn't. It is just blind to any part of reality which can't be so treated.

Choosing a coordinate system is an essential part of mathematical physics. The laws of physics are written in a form that requires the use of a coordinate system, and this cannot be removed from them. The question is which of the infinite number of equivalent coordinate systems should we choose?

This leaves us with three options:

(1) The correct expression of physical law requires that we specify a particular coordinate system.
(2) There are different expressions of physical law depending on which coordinate system we use. The mathematical formulae will be different in each coordinate system, but these systems will be entirely equivalent once we convert back to the real world. There will be rules to convert from one coordinate system to another, and when we apply these rules to the laws that govern in one coordinate system, we will find the laws we had derived in the other coordinate system. Thus every

observation or any description of physical law should only be discussed alongside a statement of which choice of coordinate system applies to that measurement or theory, so that measurements from different observers using different coordinate systems can be compared against it by transforming the results from one coordinate system to another. This places a weak constraint on the structure of physical law.

(3) The laws of physics have the same mathematical form in every coordinate system. This places a strong constraint on the structure of physical laws.

Of course, we can also combine these three options; for example by demanding that some aspects of the calculation are dependent on the choice of coordinate system, and others independent of it.

These considerations played a hugely important role in the development of twentieth century physics, both relativity and quantum physics. But that is getting ahead of ourselves; at the time of Descartes a different line of thought was pursued.

Descartes also realized that place and motion need to be treated as relative notions; there is no such thing as an absolute place or absolute rest, but only motion moving in comparison to some other observer. This meant, for example, one could explain the solar system in terms of the motion around the sun, and still accept the earth and our position as observer as central: they were just two different ways of expressing the same thing.

Descartes' mathematical achievements, and his application of them to physics, have proven to be exceptionally powerful. The rest of his philosophy, however, leaves a lot to be desired.

Having successfully combined geometry and analytic mathematics, Descartes set about trying to apply these ideas to physics. His belief (extending that of Pythagoras and the Neo-Platonic tradition) was that the real world was fundamentally geometrical – everything physical could be described by motion and extension; that being all that matter was. He felt that all knowledge about the physical world could be united in a common language, which was mathematics, and that only what can be expressed mathematically could be known for certain. That which could be treated as extension was better still, since this could be imagined as well as analysed. This, of course, implies that the knowledge supplied by our senses cannot be certain. Everything physical should be reduced to either number, dimension (which ensures that no part of reality can be spared mathematical reduction), or extension (which relates to how things are spread out and ordered). Even mass or velocity he treated as another dimension, similar to the three we are familiar with, and thus it could be fit into his system. Descartes indicated two steps in gaining understanding: intuition, which means knowledge which we do not doubt, and deduction, meaning necessary inferences from these basic ideas. His intuitive facts were that at least one self conscious thinking being exists, and various geometric rules such as triangles have three sides or Euclid's axioms.

He believed, as a prelude to the mechanistic picture, that God had started the universe by putting things in motion, and then maintained it by a general concourse. This implied that motion was just as natural to bodies as rest; that the statement should be that bodies continue in constant motion unless acted on by some other agent just as natural as saying they should be at rest unless acted upon by another agent. This is, of course, the principle of inertia and Newton's first law. Since creation, the world of extended bodies has been determined, and there is no spontaneity: everything continues to move in accordance with the principles of extension and motion. God is relegated to the first cause of the motion, the happenings of the universe continuing eternally; the world is pictured as material rather than spiritual and teleological. The only way in which motion could change was through direct contact (something which Newton contradicted). He specifically denied Galileo's forces, and did not adopt Galileo's precise definition of motion, but instead treated it as some vague concept.

Surrounding the real matter in Descartes' system was a sea of prime matter, which is invisible to us but acted to create the motions (for example the orbits of the planets) we

observe. Various vortexes and whirlpools in this invisible medium caused the orbits of the planets. While wrong, this did represent one of the first serious attempts to explain the motion of the planets without invoking final causes: it showed the world that there were other options out there.

For all Descartes bluster, he recognised that what he was proposing was only a conjecture, which needed demonstration from some reliable first principle. But what constitutes a perfectly reliable first principle? Descartes found it difficult to say, and after eliminating everything else through a method of doubt, settled on self-awareness as the only possible thing that remained and we could be certain of. From here he derived the existence and perfection of God, and from there began to build the system of his mature years.

Descartes' method of extreme doubt has its problems; it contradicts the Aristotelians, for example, who believed that knowledge began with the senses. We have two tools available to us, reason and observation, and Descartes eliminated sensual observation as a possible basis for knowledge. The main problem with this is that it leaves us with too little to go on: reason needs its premises, and these premises have to come from the senses. But surely we know that our senses are sometimes unreliable? Yes, but sometimes unreliable does not mean that they cannot be trusted at all. Firstly, we have to distinguish between accuracy and imprecision: many examples cited concerning the unreliability of our senses relate to precision rather than accuracy. Secondly, we have to be conscious of coordinate-dependent observations, and state any measurement conventions or choices alongside the actual measurement so two different observer's measurements can be accurately compared against each other. Thirdly, there are occasions when our senses do let us down. For example, take a circular disc, and divide it into three equal segments; paint one segment red, another green, and a third blue. While the disc is stationary, we can clearly distinguish the three colours. However, spin the disc fast enough, and it will appear to be a uniform white. The reason for this is that the cones in our retinas which detect colour take a short time to recharge after absorbing a photon. Spin the disc fast enough and it will rotate fully during this recharge time; the brain will detect that all three colour receptors are 'simultaneously' activated (since the eye can't measure time differences smaller than the resolution of this recharge time), and will thus conclude that the object is white. Most of the time, that would be the correct supposition to make; but in a few cases, such as this, the brain gets it wrong. So it is mistaken to say that our senses are wholly reliable. But it is equally mistaken to say on account of phenomena such as this that we cannot assume any confidence in our senses, as Descartes did. Instead, we should be aware that our senses can let us down in certain circumstances, but only in those circumstances. Once we know what those circumstances are, and the reasons why our senses break down then, we can compensate for them, either by excluding them from our sample, or adding them to the possible causes of imprecision. That such circumstances are so rare just indicates how reliable our senses usually are, and how much we can trust them in standard situations.

Descartes is perhaps best known today for his dualism. For he believed that everything physical must reduce to extension, but that the only thing we can be sure of is the mental; and mind itself is not a matter of extension. Descartes therefore put mind on a separate plane to the physical, with no connection between the two. So there are two separate worlds: the world of extended physical objects, and the world of unextended mental spirits. The trustworthiness of our understanding of sensual data must not rest on observation at all. We perceive no object by sense alone, but only by reason exercised on sensible objects. Sensations are just confused thoughts. The sensed world is vague; reason from first principles can only shed light on it. Everything that could be handled mathematically, the primary objects of geometrical bodies, Descartes placed into the world of extension. Everything that could not be, such as thought, imagination, opinions about the theatre, pain, colour, and so on is placed in the world of the mind. These secondary quantities are therefore illusions. This division was a natural consequence of Descartes' views about only the mathematical

being real. This dualism led to one of the pressing issues of modern philosophy: how the world of the mind is connected to the physical world, the mind/body problem. Why should the world of physical objects be so readily describable by mathematical forms which were only accessible to the mind? This, of course, is not an issue in an Aristotelian system, where the rational element of man is held within the form, which is naturally united with the matter in the same way as in non-material substances. However, Descartes could only resolve this by appealing to God.

To Descartes, mind was incapable of being expressed in terms of geometry or dimension. However, it is clear that our minds are related to our own bodies. There was therefore some real localisation of mind. The role of mind in the universe was small, and this is central to the whole Cartesian project, which states that the world of extension exists wholly independently of mind. This statement of the existence of mind runs against the flow of Descartes' thought, which is to express everything geometrically. Thus it was difficult for Descartes' followers to conceive that there could be anything localised but non-spatial. What Descartes meant, though, was that the mental substance (which cannot be limited to time and place) came into contact with the world of extension at one particular location. It is not therefore mind itself which is localised, but only where mind interacts with the world of extension. Of course, this still leaves many questions and difficulties.

The issue of the mind-body problem is one that has not really been satisfactorily solved in modern philosophy. Of course, there is also the issue of which parts of nature should be treated as part of the mental world, and which as part of the world of extension. For example, Descartes took colour as being part of the mental world, whereas, with science having developed, we now know it can be treated mathematically.

Descartes' own dualism is unfashionable, and perhaps Kant's division between the world of perception and the world as it actually is is more appealing to the modern ways of thinking; Kant, of course, still has the problem of how the two are related (and many followers of Kant would put the mathematical description of the world as part of the world of perception, just compounding the issue of how it applies to the real world). There is, of course, the easy path of denying that one aspect of the world exists.

Idealists would deny the real physical world, but I take this to be untenable since it completely fails to match with our experience: one has to explain why there seems to be a real world out there which obeys consistent rules and many facts which we can only learn about and later see the connections, and the best explanation of this is surely that there actually is a real and objective world out there.

At the other extreme are those, who I will call reductionists (of whom Hobbes is a good example), who would say that the physical world is all there is, and deny the existence of the mental world. Their own mental processes can, in other words, be reduced to the mathematical workings of the various particles and processes that take place in the brain. This view, however, is often viewed as being unsatisfactory because if our understanding is just a matter of mechanical processes, then it is not understanding at all. Every thought we have is really just an electrical current (or something similar), and therefore cannot be intrinsically true or false, or carry any meaning. This includes the idea that there is no world beyond a mechanical physical world; if the thought is correct then it is also meaningless, which entails a contradiction since 'correct' cannot be reconciled to 'meaningless'. Thus this viewpoint is claimed to be ultimately self-defeating. I am not completely convinced by this claim: it strikes me that it is a reasonable criticism if we adopt Cartesian mechanism, but there might be other forms of mechanism where it might not prevail.

Empiricists will affirm the real world, but say that we cannot know it: all we can know are our immediate sensual inputs, with everything else being a matter of interpretation of those inputs and thus up for doubt and questioning. This, however, does not deal with the method of falsification: we can know a great deal of what isn't correct about the real world, and by inverting this we can also gain information about what is true, albeit not

with perfect precision. Descartes would also question the empiricist assumption that our senses are more reliable than our reasoning; and I would seriously question the empiricist confusion between properties and essence (that what something is can be fully characterised by the properties which we observe; rather we should say that properties can be derived from knowledge of the essence).

Gödel's theorem, which states that each mathematical system must contain at least one axiom that cannot be proved within this system, also seems to apply here (Penrose, 1999). If we reduce everything, including our minds, to a mathematical system, then that system cannot be complete; there must be something outside it providing that fundamental axiom. However, this contradicts the claim that the mathematical conception of the universe describes everything. To avoid Gödel's theorem, there must be something in the world that cannot be translated into mathematics or expressed as a proposition. This is what Descartes suggested with his world of mind in addition to the world of extension; thus to say that everything is part of the mathematical world of extension (in whatever way we want to update Descartes' terminology and define this) is impossible. If physics (in the modern sense) can be defined as the mathematical description of nature, then there must be something out there which cannot be reduced to physics. I find the reductionist view as more believable than the other modern perspectives, but it is still not satisfactory.

One thing, however, that strikes me is a clear disconnect between Descartes' belief that matter was ultimately geometrical and Aristotle's system. Aristotle's division was between form as the part of a things essence which can be abstracted, and matter as that part of a thing's essence which can't, but instead gives it physicality. Descartes claimed that matter was ultimately geometrical; and geometry is something we can comprehend in an abstract way, as Descartes' own mathematical achievements proved. It is often said that Descartes removed the concept of form from philosophy; I am not so sure that he did not rather remove the concept of matter.

The Aristotelian view of the mind is that it is related to the form of the rational animal, which includes (but is not limited to) the mathematical description of the microscopic workings of the brain. This form is then united to matter in the usual way. Of the various modern views on the problem of mind, this is probably closest to the reductionist view, since our thought processes are clearly linked to our brain patterns. Whether neuroscience ultimately discovers about the link between thought and the brain, Aristotle's metaphysics (or rather a more mathematical version of it) can handle it. The reductionist view of the world requires that thought can be wholly expressed in terms of a one to one correspondence with brain processes and structures, but if this is found to be true it does not demonstrate that mechanism is right, because these features would be wholly described as part of the Aristotelian form. On the other hand, if it is found that brain structure alone is not enough to explain our consciousness (as Gödel seems to imply), then an Aristotelian would still be content, attributing these aspects of the brain to either the matter, the bond between matter and form, or possibly some non-mathematical aspect of the form; it is this element of a being's substance which would provide Gödel's unprovable axiom. Obviously, I am keeping this all rather vague, which is necessary since my knowledge of the underlying neuroscience is too sparse. Aristotle's system, however, differs from the reductionist view because his physics is not mechanical but contains an intrinsic directionality; the brain is not just a meaningless set of electrical impulses, but the individual final causes of those impulses combine to point at something, and from that directionality we can derive purpose, and from that purpose meaning.

However, the effect of Descartes' acceptance of mechanism on most fields has been profound. Mechanism was seen as a rival to teleology as an explanatory principle, and each success of mechanism in physics as another nail in the coffin of teleology. Thus almost every one of the humanities and softer sciences, trying to latch onto the success of physics, has largely eliminated teleology from its methodology. In this way, I think the social sciences

are even more opposed to teleology than the harder sciences; any attempt to explain social phenomena in terms of purposes and goals will leave you shunned and ridiculed. Of course, such an analogy is flawed. Even if the phenomena studied by physics are not teleological, this does not mean that it could not be useful in discussing other matters.

Outside physics, teleology was probably strongest in biology. It was teleological principles that inspired the (Western) discovery of the circulation of blood and the success of the earliest anatomists. It was design principles that inspired the understanding of the diversity of life, and how everything fit into the ecosystem. Then, of course, Darwin came along and reduced biology to underlying physical principles, and physics was supposed to be mechanical. Since then, biologists have regarded teleological explanations as pre-Darwinian superstition at least when they try to pose as philosophers. It is very difficult to discuss much of anatomy without explaining the benefit to the organism and function of various parts of the body: and *benefit* and *function* are concepts grounded in teleology. Attempts to remove such language from biological explanation have not really been successful, but that has not stopped people from trying and asserting that biology is ultimately non-teleological. There is thus in almost all fields of study a strong presumption against teleological explanations. But this is derived from the belief that physics is mechanistic, mechanism is opposed to teleology, and all other subjects can ultimately be reduced from physics. If any one of these assumptions is challenged, then the presumption against teleological explanations in other fields should be questioned.

6.9. The Legacy of Descartes

So what assessment should we make of Descartes' philosophy? His most important contribution was his revolution in how we think about space and time, including the mathematical advances of his analytic geometry. This has survived through to our latest models of physics, and I cannot see it going away. The understanding of space and time is a presumption of physics, and therefore part of metaphysics. So while the astronomical revolution and the growing understanding of kinematics in the late medieval and early renaissance periods firmly showed Aristotle's physics to be wrong, the Aristotelian could still cling to his metaphysics despite this, since Aristotelian metaphysics does not depend on his physics. But in his understanding of space and time Descartes dealt a blow to the Scholastic Metaphysics which I think still stands to this day, even though we have further developed what Descartes started. In this respect, Aristotle's system is wrong, and it needs to be modified if it is to be represented for contemporary society. However, this only entails a modification rather than a complete reversal, since space and time are not so significant in Aristotle's thought, and little else within the system depends on them, so the system can survive the original concepts being taken out, and something new put in its place and given a lot more prominence.

The second major triumph of Descartes, though shared with Galileo and Kepler, was his demonstration that the Merton calculators were right: mathematics is the key to understanding physics. Again, this has survived, indeed been strengthened, by subsequent advances in physics, and I cannot see future accomplishments undermining it. This is another blow to Aristotle's metaphysics, but one which can be survived. It means that the Aristotelian concept of form needs to be more carefully defined and at least partly mathematicised. The vagueness of the scholastics is no longer acceptable. Of course, that does not mean that the scholastics were wrong to be vague: too little physics was known in that period, and it is better to be imprecise than precisely wrong, as any guess at that time almost certainly would have been.

But there are other aspects of Descartes' thought which, if accepted, would mean a death blow to scholasticism. The first is his belief that everything can, and must, be derived from the *Cogito ergo sum*. The scholastics firmly accepted that our first point of contact with knowledge had to be our senses; it is from these we learn and understand the basic

geometrical and logical concepts that we can later make more abstract and develop into a metaphysics. The foundations of the Aristotelian system are general observations (not the specific observations of a modern scientific experiment); Descartes tried to undermine those foundations. However, this is not such a blow to Aristotle, because this aspect of Descartes' thought is rejected by almost everyone.

Secondly, we have Descartes' denial of the form/matter distinction, by saying that the real world is ultimately just extended objects, and his dualism between the world of extension and the world of cognition. However, Descartes' views here are again controversial (and were particularly so when the atomic picture was proven), and I see no reason why the Aristotelian system (properly modified) should not work as well or perhaps even better.

Finally, we have Descartes' mechanistic view of the universe; that the motions of the geometrical objects can be described by fixed and deterministic mathematical laws that exist independently of God. This deterministic view undermines the Aristotelian understanding of efficient and final causality; either they are reduced to being roughly the same thing, or perhaps we can even eliminate the concept of cause altogether, since, if this view is correct, (and moving a hundred years ahead of Descartes in the way that I am expressing it) motion can be fully expressed by solving the appropriate differential equations needing to mention neither 'cause' nor 'effect', and everything in the physical world can be reduced to the motion of extended objects. Descartes himself, of course, did not know what those laws were, and, apart from the principle of inertia and his geometry, made no real contribution to finding them. But the laws were (seemingly) found by Galileo and Newton.

This does not, however, strike a fatal blow against Aristotle's metaphysics. The premises are 1) Mechanism is inconsistent with the Aristotelian system; 2) Mechanism implies that physics can be described by deterministic laws of motion; 3) Physics is (seemingly) consistent with deterministic laws of motion. This, however, does not disprove the Aristotelian Metaphysics; that could only be done if it could be shown that Aristotle's metaphysics is inconsistent with our observations of physics. There is no argument here that Aristotle could not do just as well with the new physics as Descartes did. But it does put pressure on the Aristotelians.

The first Aristotelian attempt to derive a physics from their metaphysical principles (namely Aristotle's own work) failed rather spectacularly. That's survivable, since their physics did not only depend on their Metaphysics but brought in other assumptions: it could be those assumptions which were wrong. But the mechanists came along at the same time as Aristotle was shown to have failed, and (seemingly) got it right first time. It is then up to the Aristotelians to show that their assumptions could also lead to the same conclusions, but this was not done by the remaining scholastic philosophers, who instead continued to distance themselves from the new sciences. This does not prove that it cannot be done: they did not even try. And as the science of the mechanists gained triumph after triumph as more phenomena were brought into their system during the eighteenth and nineteenth centuries, the Aristotelians were further marginalised and pushed out of the academic world, baring a few last-ditch devotees in Roman Catholic seminaries. This does not mean that the God versus naturalism debate was settled; far from it, as there have been advocates on both sides in the mechanistic movement to the present day; but it was conducted in a language other than that of the classical theism (whether Platonic or Aristotelian) in which Christianity was first and is best expressed. It was from this stream that the abomination of liberal Christianity was born, and also the Neo-orthodox Protestant representations. The old versions of Christianity, of course continued (though in some cases diluted with the more modern streams) and have endured to the present day, whether Thomist Roman Catholicism, Platonic Eastern Orthodoxy, or classical reformed Protestantism (which, where it is philosophical, has roots in a softened version of Aristotle which suppresses much of the speculation of the Romans), but were shunned by the scientists.

Of course, the modern philosophers since Descartes either assumed this mechanistic picture (as the likes of Spinoza did), or accepted but added to it (such as Kant), or reacted against it (such as in the relativists or existentialists). While few people today directly follow Descartes, in the science (as opposed to philosophy) departments of the universities, the idea that there is nothing beyond the mechanistic view of physics is also quite common. The relativists and existentialists I dismiss as being ultimately incoherent and anti-science; they focus on individual experience, and thus have no means to access the objective truths that science shows to us, or to explain why science is so successful.

Kant's *Criticism of Pure Reason* contained an attempt to defend Newtonian principles, including a view of Mechanism similar to Descartes, on the basis of his fundamental axioms (the thesis that Kant was defending was that 'every material *event* has a material cause'). His *Critique of the Teleological power of judgement* questions whether Mechanism alone is sufficient for an understanding of nature. In this work, Kant tried to judge between the maxims "All generation of material things and their forms must be judged as possible in accordance with merely mechanistic laws" (which is best interpreted to mean that mechanistic laws alone can account for everything) and "Some products of material nature cannot be judged as possible in accordance with merely mechanistic laws". His conclusion has been understood in two ways, either that sometimes non-mechanistic explanations (i.e. teleological) are necessary in addition to mechanistic explanations, or that sometimes non-mechanistic explanations are necessary as an alternative to mechanistic explanations. The first of these is more consistent with Kant's general philosophy, and should be preferred. Kant does not deny mechanism, but instead states that this does not always give a full description of the causes. Thus at least parts of Kant's system depend on the assumption that the mechanistic world-view is correct.

Hume is harder to characterise; although he admires Newton's achievements, his admiration is qualified. Unlike the other two giants of modern philosophy (Descartes and Kant), or the giants of early modern physics, Hume did not have a good mathematical education, and perhaps did not realise the importance of mathematics in the new science. This is a major limitation of his thought and a source of many of his errors. Like Descartes and Kant, however, Hume only considered event causality, and it was this that he attacked. His purely empirical views on causality led to the conclusion that anything that could be imagined was in principle possible, which lies against the views of Descartes and the later mechanists. Like Galileo, he was committed to empirical investigation as a first principle, unlike Galileo he did not go beyond this into mathematics and theory. He dismissed talk of both Aristotelian substance and Galileo/Newton's forces for the same reason: that they were trying to go beyond our direct impressions, and so were just fictions put in by the human mind to try to explain what we saw but which need not have any bearing on the actual world; he was, of course, more dismissive of *substance* than *force*. When the sciences talk about forces or powers, these words must be reinterpreted.

> There are no ideas, which occur in metaphysics, more obscure and uncertain, than those of power, force, energy or necessary connexion, of which it is every moment necessary for us to treat in all our disquisitions ....
>
> In reality, there is no part of matter, that does ever, by its sensible qualities, discover any power or energy, or give us ground to imagine, that it could produce any thing, or be followed by any other object, which we could denominate its effect. Solidity, extension, motion; these qualities are all complete in themselves, and never point out any other event which may result from them. The scenes of the universe are continually shifting, and one object follows another in an uninterrupted succession; but the power of force, which actuates the whole machine, is entirely concealed from us, and never discovers itself in any of the sensible qualities of body....

> Since, therefore, external objects as they appear to the senses, give us no idea of power or necessary connexion, by their operation in particular instances, let us see, whether this idea be derived from reflection on the operations of our own minds, and be copied from any internal impression. It may be said, that we are every moment conscious of internal power; while we feel, that, by the simple command of our will, we can move the organs of our body, or direct the faculties of our mind. An act of volition produces motion in our limbs, or raises a new idea in our imagination. This influence of the will we know by consciousness. Hence we acquire the idea of power or energy; and are certain, that we ourselves and all other intelligent beings are possessed of power. This idea, then, is an idea of reflection, since it arises from reflecting on the operations of our own mind, and on the command which is exercised by will, both over the organs of the body and faculties of the soul (Hume, 1748, 7.48-51).

For Hume, therefore, the ideas of force, energy and so on were no more than matters of our own self-reflection; convenient labels which had no real existence in nature. He seemed to be unaware of the precise definitions given by Newton (following Galileo). He views Newton's achievements as no more than a way to keep track of phenomena. He praised Newton for his arguments against simple mechanism – that force was only applied by bodies coming into contact with each other (as Newtonian gravity is action at a distance), believing that this supported his point that ultimate causes and explanations could not be known.

Hume thus stands as something of an anti-Descartes, rejecting pretty much everything that Descartes believed in (baring their common rejection of Aristotle and in particular his final causes).

We therefore see that of the major schools of Enlightenment Philosophy, Descartes and Kant depend on a deterministic mechanical view being correct (so if it is incorrect, then both of these are wrong, at least in part), the modern reductionists depend on a mechanistic view without the nuances of Descartes and Kant who both leave room for alternative explanations in addition to mechanism (so if mechanism is incorrect, then these are wholly wrong), while Hume and the more modern movements such existentialism reject not only the mechanical world view but any world view that seeks to explain the wider world (so if the world is explicable beyond our direct sensual impressions, then these thinkers are at least in part wrong).

CHAPTER 7

Classical Physics comes to Maturity

7.1. From Galileo to Newton

7.1.1. Galileo

As a young man, Galileo[1] became fascinated by the new Mathematical ideas that had emerged, and their relationship to natural philosophy. He started off life studying medicine, but found himself distracted into the study of mathematics and natural philosophy. The traditional story is that he was fascinated by the regularities of a swinging chandelier in Pisa cathedral. He switched from medicine to mathematics and natural philosophy. Appointed to a teaching position at the University of Padua, a place heavily influenced by the medieval mathematicians, he studied the Merton Calculators, Buridan, Oresme, and the rest of the scholastic mathematicians, and was captivated by their work and ideas; and their ideas formed the inspiration for his life's work. He also studied the now re-discovered writings of Philoponus. In his earlier treatises and personal notes, written as a young man, he carefully acknowledged these sources. However, in his later works, he did not, and thus the illusion was spread that most of his theoretical ideas were his own.

That is not to diminish his achievements. He selected the right material, realized its importance, polished and improved it immeasurably, and did the one thing that the medieval mathematicians did not do: tested the ideas against experiment. Much of Galilean physics was originally proposed by the likes of Buridan and Oresme several centuries earlier (and restated and improved by subsequent pre-Galilean writers, such as Domingo de Soto); Galileo not only proved them right when the rest of the world ignored them, but also improved and perfected them. Of course, his best known influence was Copernicus and Kepler's early defence of the heliocentric system. His predecessors were great men, but Galileo was a genius; and just as renaissance society was moving away from the progress made by the medieval philosophers, Galileo put Western Europe's intellectual society back on track and gave it fresh impetus.

Galileo first found fame in Italy as an exceptional mathematician, and for his invention of the thermometer and a hydrostatic balance able to measure weights to an unprecedented precision. For some time, Galileo held a teaching position at the university of Padua, and gradually developed his ideas concerning kinematics and motion. However, his international recognition arrived through his astronomical studies. The telescope was invented in the Netherlands when Galileo was roughly forty, but as soon as he learnt about it he built his own, improving on the original design. His discovery of the four major moons of Jupiter was a sensation, showing firstly that the solar system was more complicated than had believed, but also that there were astronomical bodies which rotated around Jupiter rather than the earth. His observation of the phases of Venus was perhaps more significant. What he observed was contrary to what was expected by Ptolemy's system, but consistent with

[1]My discussion here is primarily adapted from the achievements of Edwin Burtt (Burtt, 1931); my only criticism of Burtt's work is that he (mistakenly) did not consider the work of the Medieval writers I have previously discussed significant – at least he made no mention of them and thus attributed to Galileo some ideas which were not original to him. However, baring this proviso, I do not think that Burtt's work has been bettered. (Hannam, 2010) is a good guide to the historical context of Galileo and in particular his disputes with the Roman Catholic Church and inquisition.

Copernicus and Tycho. This was the observation that dealt the fatal blow to Ptolemy's astronomy.

Galileo was a firm advocate of the Copernican system, and this, as is well known, brought him into trouble with the Church. The reasons why the geocentric system were preferred were partly scientific and partly theological. Galileo's opponents were quick to seize on the scientific arguments against his system, particularly the parallax which Galileo could never answer satisfactorily. Galileo did not have a scientific case, so he tried to make a theological case that the sun was at the centre of the solar system. Perhaps at other times he might have got away with this, but not while the effects of the reformation were being felt in Rome.

For Rome, the central issue with the reformation was that people like Luther and Calvin claimed the authority to interpret the Bible for themselves, and that the Roman Church (as well as the Eastern Orthodox Church) had got several of its interpretations wrong, and was thus capable of misinterpreting scripture. The Bishop of Rome and his Cardinals saw this as a direct attack on their authority (not surprisingly, because it was). They had also made interpretations of scripture in defence of geocentricism, ignoring Augustine's warning (Augustine of Hippo, 404-420) that the Christians should be careful when stepping outside the basic historical and theological foundations of Christianity, lest their opponents, finding them ignorant on matters (such as concerning what we now call science) about which the Bible was ambiguous but those opponents expert, should equally view the Christians ignorant about the core doctrines of Christianity (where in reality the Christians would be expert and the opponents ignorant). To be fair to the Church authorities, that the earth was the centre of the universe seemed as scientifically certain as anything could be, so there was a natural temptation to claim bragging rights and say that the Bible had got there first. Galileo, on the other hand, said that those Biblical verses should be taken allegorically, but other Christian principles could be used to support heliocentricism. This, of course, was interpreting the Bible for himself against the approved teachings of the Church – the very same 'error' that the reformers had made. Thus Galileo was given the same treatment. And to make clear that he was in error, the Roman Church ruled in 1617 that the Sun went round the earth, and not the other way round.

This, of course, proved highly amusing to the Protestants, particularly when Galileo (or rather Kepler's similar ideas) was 'proven' right a little over half a century later. The man given over to the suppression of Galileo's works was Robert Bellarmine, one of the most loyal and greatest intellects the Roman Church had available to it at the time. He was a leading figure in the battle against Protestantism. Anyone who opposed Bellarmine was seen as an ally in the fight against Rome, so the Protestants, unfairly to Galileo who remained to his dying day a devout Catholic in all other matters, claimed him as one of their own. The treatment of Galileo should not, however, cause concern about Christianity, except perhaps the most extreme part of Roman Catholic Christianity. It was not a battle of science against religion, but one set of religious arguments against another: the science at the time was, if anything, on the Church's side. It should always be a reminder to the Church to remain neutral about matters concerning which it has no expertise (even those matters which appear scientific certainties).

Of course, Galileo was spared from the full wrath of the inquisition by a secular authority who gave him shelter. This gave Galileo time to write up formally and present his ideas, and his best known works of this period are his dialogues. Dialogues, if done well, are a good way of handling controversial topics in a time of official persecution. Few people have done the philosophical dialogue well, since it requires literary as well as philosophical skill. Plato was, of course, the master of this form, and has never been matched. Among the major modern philosophers, Hume was probably the only one who composed readable dialogues. But Galileo proved adept at presenting his ideas in this way. A dialogue has the advantage that you present at least two points of view in the mouths of different characters. They can

present their arguments and rip their opponents to pieces and defend themselves against the attacks of the opponents. It thus, if done right, gives the author a good chance to really explore the pros and cons of each position, without having to commit himself to either view. If pressed, he could always say that the voice defending the conventional view was your own, and he wrote the dialogue in defence of the conventional position rather than as an attack on it. A writer can also spread his own views among the different characters. Of course, a dialogue has the disadvantage that it is harder to cite sources than in a conventional text, and they are harder to write, since to make it believable, you have to be a master of your opponent's position as well as your own.

In addition to the dialogues which defended the Copernican system (Galilei, 1632), presented his experimental evidence for the work of the Calculators and brought the ideas (if not the names) of the pioneers of mathematical science to a wider audience (Galilei, 1638), the other major work of Galileo's maturity was his Assayer (Galilei, 1623).

Galileo's method was in three steps: intuition, demonstration, and experiment. We isolate as fully as we can a particular phenomenon, extract from it its simple and absolute constituents, which could be treated mathematically (note that he assumes that this extraction is possible; i.e. that any complex phenomena can be deconstructed into simpler parts). Having performed this step, we no longer need the observed facts. We can apply the mathematical method to our abstraction, and deduce whatever is possible or necessary. However, to check and make our results more certain, we need to develop possible demonstrations which can be verified by experiments. With these principles established, we can then use them as a basis to understand more complex phenomena.

Galileo's principal belief was that physics ought to be treated in terms of well-defined mathematical terms. He gave an exact definition of terms such as velocity, acceleration, force, and so on, whereas in scholasticism these terms had been left vague and rarely used. On the other hand, the key terms of scholastic metaphysics, form, matter, substance and so on he neglected, because they could not be given a precise mathematical definition. Once again we should take note of the reason for their neglect: he did not disprove the logic behind them, but just claimed that they were not necessary for what he was trying to do. Indeed, I consider Galileo's objection to be pertinent: any defender of the metaphysical concept of form has to elaborate on how it connects to the more mathematically expressed scientific concepts. If form, which is meant to describe physical objects, cannot connect to mathematical physics, then it is indeed useless to the physicist. This is not to say that the concept of form is useless in general, only that its defenders need to do more work. The problem for the defenders of scholasticism of Galileo's time was not so much that their philosophy was undeveloped, but that the physics of their day was not sufficiently developed to really express what form was.

Galileo believed that if one knew the correct laws and initial conditions, then one could compute everything without need of experiment. If one fired a projectile, one would not need to look where it landed because that would already have been calculated; the observation would tell us nothing new (note that this presupposes physical determinism). And, of course, Galileo believed that it was possible to know what those laws are. This distinguishes Galileo from the Empiricists.

However Galileo did not take this to the opposite extreme held by Descartes. He accepted experiment as a key step in verifying that one's understanding of nature is correct, and resolving disputes. He trusted our senses as reliable (at least under normal circumstances) and thus an important part of understanding motion. Sensible facts are there to be explained; they cannot be overruled or ignored on account of reason, and the primary purpose of philosophy is to explain what our senses reveal to us. Sensing by itself is not enough; we also need to understand. Thus Galileo differed from Descartes in setting as an a priori axiom that the senses provide us with reliable information, and with the Empiricists in saying that sensual data was not all that we could know. Indeed Galileo saw himself as

a disciple of Aristotle (though not the Aristotelians) in this regard; the difference between them being not in method but in the experimental information available. Had Aristotle seen what Galileo had seen, then he would have drawn different conclusions (Burtt, 1931). This can be contrasted with Bacon's view of Aristotle. Of course, like all these contemporaries, Galileo was dismissive of arguments from authority:

> Sarsi goes on to say that since this experience of Aristotle's has failed to convince us, many other great men also have written things of the same sort. To this I reply that if in order to refute Aristotle's statement we are obliged to represent that no other men have believed it, then nobody on earth can ever refute it, since nothing can make those who have believed it not believe it. But it is news to me that any man would actually put the testimony of writers ahead of what experience shows him. To adduce more witnesses serves no purpose, Sarsi, for we have never denied that such things have been written and believed. We did say they are false, but so far as authority is concerned yours alone is as effective as an army's in rendering the events true or false. You take your stand on the authority of many poets against our experiments. I reply that if those poets could be present at our experiments they would change their views, and without disgrace they could say they had been writing hyperbolically – or even admit they had been wrong (Galilei, 1623).

The key to understanding the world was mathematics. Physical bodies are geometrical figures. The method to resolve a problem is thus to first express it in pure mathematical terms, then to seek to understand what that mathematics implies. Sometimes this involved overriding the natural impression given to us through our senses (such as the earth is stationary). The senses alone cannot give us the rational order on which the world is founded. However, the question remains whether the mathematical world is fundamental, or whether there is a foundation behind it. Galileo, despite his struggles with the church, remained a resolved Roman Catholic. He understood that underneath the mathematical layer of reality was the mind of God. God was thus a mathematician; the world he created was mathematics made flesh (as it were). It is in God that the fundamental axioms depend. God's knowledge differed from ours only in that God knew more propositions (indeed all propositions) whereas we only know a few. That does not mean, however, that we cannot understand the few propositions we do know accurately and precisely (in this, Galileo differs from those modern sceptics and relativists who state that our knowledge is fundamentally imprecise, vague or subjective). Thus Galileo believed that the various religious texts which seemed to contradict nature must be able to be, and should be, reinterpreted in a way that is consistent with nature. Both Science and the Bible had the same author, who could not contradict himself.

The key to mathematics, on the other hand, is to make things precise and well defined. Galileo thus divided the world between primary and secondary quantities. The primary quantities are absolute, objective, immutable and mathematical; these are matters of divine and human knowledge. The secondary quantities are subjective, and fluctuating. These are matters of illusion[2]. Our senses give us hints of the true structure underlying them, which consists of number, figure, magnitude, location and motion, which can be expressed mathematically. Furthermore, secondary quantities are subjective. They are the primary quantities passed through the filter of our senses, which makes them less clear. That is

[2]Burtt, *The Metaphysical Foundations of Modern Science*, 3.C uses the word opinion here. I feel that this redefines opinion away from its customary meaning. I have the opinion that $2+2=4$, and anyone who disputes this needs revisit their mathematical education. That this opinion is also a matter of fact does not change that it is an opinion. True statements can be disputed (as O'Brien in Orwell's *1984* wanted to dispute the basic arithmetic I alluded to); therefore that something is disputed does not mean that it is really subjective; only that at least some of the people who dispute it are ill-informed and in error.

to say that our senses add distortion; we see a cannonball rather than a geometric object because of the way that our senses filter the information they receive[3]. Because the sensed quantities were only secondary, Galileo noted that the universe did not rely on our senses. The primary quantities did not depend on us. On the other hand, sounds and smells are the interpretation of those primary quantities as mediated through our senses. Thus if our ears were taken away, there would be no sounds, and if our noses would disappear then no smells, even though the motions we interpret as sounds and smells remain. Thus man is no longer primary to nature; nature is not our experiences, but our experience is secondary to nature – we are just a spectator. The world does not revolve around or depend on man. The primary quantities, the underlying reality, is not dependant on man. To us (baring a few philosophical – both of the modern and ancient varieties – extremists) this conclusion is natural; in Galileo's day it was a revolutionary. The natural world thus has an existence independent of both man and God.

> Now I say that whenever I conceive any material or corporeal substance, I immediately feel the need to think of it as bounded, and as having this or that shape; as being large or small in relation to other things, and in some specific place at any given time; as being in motion or at rest; as touching or not touching some other body; and as being one in number, or few, or many. From these conditions I cannot separate such a substance by any stretch of my imagination. But that it must be white or red, bitter or sweet, noisy or silent, and of sweet or foul odour, my mind does not feel compelled to bring in as necessary accompaniments. Without the senses as our guides, reason or imagination unaided would probably never arrive at qualities like these. Hence I think that tastes, odours, colours, and so on are no more than mere names so far as the object in which we place them is concerned, and that they reside only in the consciousness. Hence if the living creature were removed, all these qualities would be wiped away and annihilated. But since we have imposed upon them special names, distinct from those of the other and real qualities mentioned previously, we wish to believe that they really exist as actually different from those.
>
> Then there remains the air itself, an element available for sounds, which come to us indifferently from below, above, and all sides-for we reside in the air and its movements displace it equally in all directions. The location of the ear is most fittingly accommodated to all positions in space. Sounds are made and heard by us when the air without any special property of "sonority" or "transonority" - is ruffled by a rapid tremor into very minute waves and moves certain cartilages of a tympanum in our ear. External means capable of thus ruffling the air are very numerous, but for the most part they may be reduced to the trembling of some body which pushes the air and disturbs it. Waves are propagated very rapidly in this way, and high tones are produced by frequent waves and low tones by sparse ones.
>
> To excite in us tastes, odours, and sounds I believe that nothing is required in external bodies except shapes, numbers, and slow or rapid movements. I think that if ears, tongues, and noses were removed, shapes and numbers and motions would remain, but not odours or tastes or sounds. The latter, I believe, are nothing more than names when separated from living beings, just as tickling and titillation are nothing but names in the absence of such things as noses and armpits. And as these four senses

[3]This idea seems to owe something to Plato.

> are related to the four elements, so I believe that vision, the sense eminent above all others in the proportion of the finite to the infinite, the temporal to the instantaneous, the quantitative to the indivisible, the illuminated to the obscure–that vision, I say, is related to light itself. But of this sensation and the things pertaining to it I pretend to understand but little; and since even a long time would not suffice to explain that trifle, or even to hint at an explanation, I pass this over in silence.
>
> Since the presence of fire-corpuscles alone does not suffice to excite heat, but their motion is needed also, it seems to me that one may very reasonably say that motion is the cause of heat. But I hold it to be silly to accept that proposition in the ordinary way, as if a stone or piece of iron or a stick must heat up when moved. The rubbing together and friction of two hard bodies, either by resolving their parts into very subtle flying particles or by opening an exit for the tiny fire-corpuscles within, ultimately sets these in motion; and when they meet our bodies and penetrate them, our conscious mind feels those pleasant or unpleasant sensations which we have named heat, burning, and scalding. And perhaps when such attrition stops at or is confined to the smallest quanta, their motion is temporal and their action calorific only; but when their ultimate and highest resolution into truly indivisible atoms is arrived at, light is created (Galilei, 1623).

Galileo also adopted the atomic theory of matter, seeing that if all of nature could be reduced to small indivisible atoms[4], then he could explain the transmutation of substances without recourse to changing forms. These atoms are just mathematical quantities, and their motions indirectly lead to our sensual experiences.

Another way in which Galileo differed from the Scholastics was in his discussion of causality. For the scholastics, causes were substances (at least for efficient and final causes; formal and material causality are a little different), so one substance is the cause of another substance. For Galileo, the causes were *forces*, which are not substantial. So, for example, in the scholastic system gravity was explained by saying that a brick suspended in mid air had a final cause of being a brick lying on the ground a short time later (that is to say that things made primarily of earth have the natural tendency to move downwards); Galileo described gravity as being a force that acted on everything giving constant and uniform acceleration towards the surface of the earth. The reason gravity is there is not something internal to the body (a final cause) but external to it (the force). However, Galileo was unable to answer the question of what a force fundamentally was; he left the matter open.

Kepler had already replaced the formal cause with the geometrical harmonics of nature; however Galileo was not so interested in form as with the dynamic universe. He was not immediately interested in the purpose or use of something, but only the immediate question of how things moved. Purposes could be dealt with later; it was better to understand *how* and *what* before asking *why* (which could be explained mainly in terms of its use to man). Final causality was banished from consideration of the primary workings of the universe, because purpose (which is, of course something different from final causality, but again this distinction was ignored) was a matter of the secondary qualities related to man. This of course left God in a difficult place, because traditionally conceived God was the ultimate provider of purpose; but if that were so it would cut God off from the foundational reality of the universe. However, God could still be thought of as the ultimate efficient cause, namely the one who started the universe in motion. This seemed to fit in with the Christian idea

[4]Note that what we refer to as atoms are not the same thing as what Galileo referred to as atoms; Galileo would be referring to either quarks, leptons, and gauge Bosons if these are indeed fundamental; or the more fundamental objects otherwise. Somebody in the early nineteenth century was a little over-eager when they applied the name to the constituents of the chemical reactions they studied.

of God as creator, although in reality it leads to a more deistic view of the universe. There could only be one true and primary cause of similar effects, and this cause and its effects were continuously connected. For every type of motion, there is some force which can always be depended on to create its effects. These causes are to be identified by identity, uniformity and simplicity; and in particular the whole system was deterministic: cause always led to effect, and effect was always proceeded by cause. This idea of necessary connection paved the way for the notion that cause and effect are really the same thing, just two different points along the solution of an equation. Thus Galileo's system leads invariably to a mechanistic world view. Purpose was banished from physics, and final causality from philosophy.

However, in my view Galileo's most important philosophical achievement was in his discussion of space and time, and he (along with Descartes) revolutionised the concepts compared to the prevalent Aristotelian perspective. So complete was the revolution that to me the Galilean way of thinking seems entirely natural, and Aristotle's unnatural. However, to understand Galileo's genius we should not forget that in his time the reverse was true. However, I shall discuss space and time in its own section.

Galileo is rightly regarded as one of the foremost and most important intellects in human history, and as the founder of modern science. Of course, he was not completely accepted at his time, and for some years his ideas remained controversial. But his influence spread, and if anything his struggles against the Roman Church's hierarchy only increased his appeal and influence in Protestant Britain, where his ideas would be completed. His importance has been overstated in some senses, to the detriments of his predecessors. Many (not all) of his ideas had been previously suggested by the Medieval mathematicians and ancient philosophers, and his empirical studies had also been foreshadowed in the medieval and renaissance periods. But Galileo was the one who merged the two streams, and advanced them from ideas on a page to a full and complete system, and in particular extended the mathematical methods that had long been used in Astronomy to kinematics on the earth. Bradwardine, Buridan, Oresme, Copernicus, Nicholas of Cuda, Roger Bacon, Dominic de Soto and Philoponus were all great men, but Galileo was the genius that the emerging science needed. The tide of the times was turning against these ideas, and back towards a 'pure' Greek and Roman view of physics. Galileo, along with his contemporaries Kepler and Descartes, permanently reversed that trend. The modern world was born.

7.1.2. Newton

Like many of the giants whose shoulders he stood on, Newton was first and foremost a brilliant mathematician. His accomplishments in physics and the other fields he was interested in have almost entirely been superseded (at least once we go down to the details); his mathematics is still at the centre of how we think about physics. However, unlike Descartes and the Philosophers, he did not believe that the universe itself was mathematical. Mathematics could be used to describe certain aspects of the universe, which once extracted, could then be considered using that tool. However, mathematics was not everything, and one could not uncover the universe by mathematical reasoning alone. Instead, Newton believed that experimentation was the key to understanding natural philosophy. The results of these observations were to be understood in mathematical terms, and conclusions drawn from them and mapped back to nature, but the fundamental building blocks of the universe were not geometrical forms, but something different. Newton thus found the correct balance between the mathematical and empirical strands of his predecessors. With the possible exceptions of Galileo and Roger Bacon, the earlier thinkers had strayed too far towards either the mathematical or the empirical extremes. Newton got it right.

The other major difference between Newton and his predecessors is that he he represented a transitional stage between the geometrical-centred approach of the early modern physicists and the algebraic approach favoured by later physicists. Descartes had, of course, already shown how to translate geometrical objects into algebraic symbols, but he himself

did not realised the implications of this and still believed that nature was basically geometry. Newton still believed that geometry was fundamental, and continually went back to the geometrical ideas and tried to think about things geometrically wherever he could. However, by using algebra where geometry was insufficient, Newton was able to approach the problems in a different and more successful way. If anything, Newton's devotion to the geometrical principles of his predecessors was a significant limitation; his version of calculus, for example, used a far more cumbersome notation than that of Leibniz, and it is Leibniz's formulation of calculus that we continue to use today rather than Newton's.

Newton's goal in his major work was to outline the mathematical principles of natural philosophy. He uses the tools that he develops to derive various propositions linking the sun, planets, moons and tides; while expressing confidence that the rest of the phenomena of nature could be derived by considering either attractive or repulsive forces by mechanical principles. He is aware that his work might not be the final word, but hopes that his methods may shape some light on the correct principles.

The emphasis on *method* is important. Newton was not much of a philosopher; his principal goal was to come up with a way in which physical phenomena could be investigated and understood. His work was not only mathematical, but also experimental, particularly in his studies of optics and the phenomena of reflection, refraction and the mixing of colours. Of course, his ideas concerning optics - that light was fundamentally a particle, denying its wave-like behaviour – were rejected, but his experiments were certainly important, as was his extension of the mathematical method to the study of colour. For Newton, far more than his mathematical predecessors (again, with the possible exception of Galileo), experimental guidance must accompany every step of the explanation of phenomena. There was no *a priori* certainty that the world was through and through mathematical. This was something that could only be seen through testing, and only to the extent to which experiment allowed. If we could discover mathematical laws, then great. If not, then we would have to either extend our understanding of mathematics or use some other tool to understand nature. We should not assume that the universe is mathematical, nor claim it to be so more than our experiments suggest. We should also not think that the mathematical description covers all of reality, nor reduce reality to only that which can be described mathematically, since neither of these could in principle be shown by experiment.

Newton recognised the difference between a mathematical model or even a mathematical description and reality itself. For this reason, he was adamant that every conclusion drawn from the mathematical description had to be tested experimentally: it could be at that point where the description of nature broke down. Equally, mathematics ultimately came from mechanical principles: it is while studying the universe that we are inspired to understand the ideas of geometry; by seeing approximate circles in nature, we are inspired to consider the properties of the exact circle. By counting natural objects, we are inspired to consider abstract arithmetic.

Of course, Newton had little patience for the scholastic notions of substantial form, and like all of the early modern physicists he did not consider final causality. However, he was critical of the attempts to remove ideas of purpose from the universe. Boyle himself, one of the most important influences on Newton, had made it clear that while many of the purposes underlying the universe could not be discovered, we should not give up hope of finding some of them (Barrow and Tipler, 1986). For example, Boyle stressed the well known analogy involving the watch and watchmaker. Newton, being an opponent of Descartes' physics as well as his metaphysics, was also inspired to reflect on the underlying purpose of the universe.

> Did blind chance know that there was light, and what was refraction, and fit the eyes of all creatures after the most curious manner to make use of it? These and other suchlike considerations, always have, and always will prevail with mankind, to believe that there is a Being who made

> all things, who has all things in his power, and who is therefore to be feared. (Newton, 1950)

Newton was thus in agreement with the underlying design arguments. One of his correspondents, Richard Bentley, gave a detailed argument based on Newtonian mechanics about how the underlying mathematical order showed evidence for an intellectual composer of that order; the rationality of God shown forth in the sublime mathematical principles.

> Bentley was to argue that design is most clearly witnessed by elegant mathematical laws of a general and invariant character rather than the specific, but relative, adaptations we see in the animal world. He attempted to construct a eutaxiological design argument based upon our knowledge rather than, as had often been the case, a teleological argument founded on our ignorance (Barrow and Tipler, 1986).

Thus in Newton's own country there was a strong movement to use Newton's discoveries to support the design arguments, and indeed build new and better design arguments; not founded on God-of-the-gaps ignorance (which has always been and continues to be a very sloppy method of argumentation), but rather based on and derived from scientific knowledge. Newton himself was generally supportive of such efforts. This group of Newton-inspired thinkers was possibly Hume's main target in his *Dialogues Concerning Natural Religion.* But of course, Newtonian design was different from scholastic final causality. Once again, the mechanical laws of nature, external to the particles they governed, stood as an intermediary between God and the universe.

Newton was a deeply religious man. He might have opposed the scholastic understanding of physics, but he was fully with them in supporting the idea that God was ultimately behind all the laws that he was uncovering, and that all physical phenomena led one back to an understanding and awe of God,

> This most beautiful System of the Sun, Planets, and Comets, could only proceed from the counsel and dominion of an intelligent and powerful being. And if the fixed Stars are the centres of other like systems, these being formed by the like wise counsel, must be all subject to the dominion of One; especially since the light of the fixed Stars is of the same nature with the light of the Sun, and from every system light passes into all the other systems. And lest the systems of the fixed Stars should, by their gravity, fall on each other mutually, he hath placed those Systems at immense distances from one another.
>
> This Being governs all things, not as the soul of the world, but as Lord over all: And on account of his dominion he is wont to be called Lord God, or Universal Ruler. For God is a relative word, and has a respect to servants; and Deity is the dominion of God, not over his own body, as those imagine who fancy God to be the soul of the world, but over servants. The supreme God is a Being eternal, infinite, absolutely perfect; but a being, however perfect, without dominion, cannot be said to be Lord God; for we say, my God, your God, the God of Israel, the God of Gods, and Lord of Lords; but we do not say, my Eternal, your Eternal, the Eternal of Israel, the Eternal of Gods; we do not say, my Infinite, or my Perfect: These are titles which have no respect to servants. The word God usually a signifies Lord; but every lord is not a God. It is the dominion of a spiritual being which constitutes a God; a true, supreme, or imaginary dominion makes a true, supreme, or imaginary God. And from his true dominion it follows, that the true God is a Living, Intelligent and Powerful Being; and from his other perfections, that he is Supreme or most Perfect. He is Eternal and Infinite, Omnipotent and Omniscient; that is, his duration reaches from Eternity to Eternity; his presence from

> Infinity to Infinity; he governs all things, and knows all things that are or can be done.
>
> ...
>
> We know him only by his most wise and excellent contrivances of things, and final causes; we admire him for his perfections; but we reverence and adore him on account of his dominion. For we adore him as his servants; and a God without dominion, providence, and final causes, is nothing else but Fate and Nature. Blind metaphysical necessity, which is certainly the same always and every where, could produce no variety of things. All that diversity of natural things which we find, suited to different times and places, could arise from nothing but the ideas and will of a Being necessarily existing. (Newton, 1726)

Newton thus gives glory to God for the ordering of gravity, the laws of which Newton himself discovered. All things are under the dominion of God, and Newton's discoveries are only to uncover the mechanisms by which that dominion is exercised.

Of course, particularly on the continent of Europe (and in Scotland), this emphasis on design was opposed. The likes of Leibniz and Spinoza are prominent in this discussion. They believed that the new mechanical principles left no room for purpose. Leibniz believed that there were many possible worlds which God could have created, but He choose the one that had the maximal degree of perfection. This lead to final causality being joined to mechanical principles: they were basically just the same thing. The harmony of all things was a consequence of this being the best possible world: God created a world in harmony, and as the mechanical principles evolved the universe, that harmony remained. Leibniz thus still appeared to design in a sense (the 'best possible world' is a design principle), but in a different way to the traditional ideas. It is not so much benefit for mankind which is the underlying principle, but that nature itself should (at a fundamental level) be as harmonious as possible. The world, however, was not made for us alone. Leibniz's main opposition to the Newtonians (most prominently Samuel Clarke) was that they believed that God was constantly needed to correct aberrations in the universe and uphold the laws of motion; to Leibniz this idea was anathema: God was not needed beyond as a creator, since the laws continued independently.

Spinoza (writing before Newton) was more critical of the attempts to introduce design arguments into the new physics. He believed that such notions as design only arose because of our ignorance of the mechanical laws of nature and our gullibility. Mechanism was completely opposed to such principles; it showed that nature does not propose itself to any end in its operations, and all final causes are nothing but pure fictions of human imagination. Among the philosophers, Hume and Kant were probably most influential, and Hume certainly firmly rejected any notion of design entering into our arguments, while Kant was more ambiguous. Hume was, of course, firmly rejected by most scientists of his time because of his absurd views on causality and opposition to Newton; by Kant's time, the distance between physics and philosophy had increased, and he has had little impact on how physicists themselves think.

Thus among the early physicists, there was a considerable, if not universal, acceptance of the place of design in explanations of the world. These physicists, at least since Newton, were well aware of the limitations of their methodology and that there was no guarantee that everything could be explained through mechanical law: it was just an idea which seemed to work well for certain phenomena up to the available experimental precision. These design principles were not only seen in the biological God-of-the-gaps adaptation, but also arose from the structure of physical law itself. Among the early modern philosophers, however, there was a considerable, if not universal, rejection of design principles, with the understanding that mechanism either made them superfluous or (more frequently) denied

them. Such a state of affairs remained, I think, until Darwin and Huxley shook things up, while more and more phenomena came to be understood in terms of mechanics.

Newton's most important contribution was the discovery (although Leibniz independently uncovered it at about the same time), of calculus, the mathematics that describes change. This was the ingredient that Newton's predecessors missed. Thanks to Newton, instantaneous velocity and acceleration were given a precise mathematical definition; the Laws of Motion were reduced to various equations which could (in some cases) be easily solved.

Calculus is built around two principles, differentiation and integration. Differentiation measures the gradient of a curve. We draw a curve on a graph that represents some mathematical function, and next pick a particular point on that curve. We draw a line intersecting that point and parallel to the curve at that point. The differential of the function gives the slope of that line. The use of this is seen when we consider velocity, which is distance travelled in a given time. I start by considering a one dimensional case, where the particle is travelling in a constant direction. We plot distance travelled against time, then the curve corresponding to a particle travelling at constant velocity would be a straight line, with the velocity as its slope. But what if the velocity is constantly changing? It is a straight forward to extend the constant-velocity definition using the concept of differentiation. We draw a line parallel to the curve at a particular point, measure the slope of this curve, and that gives us the instantaneous velocity. More formally, if we write location x (just considering one dimension, i.e. straight line locomotion, for the moment) as a function of time t, (so we have $x(t)$) then we can write the velocity v as the differential of x with respect to t, which is formally written as $v = \frac{dx}{dt}$. Acceleration is treated in the same way. This time we plot velocity against time, and write the instantaneous acceleration at some particular moment of time as the gradient of that curve at that moment of time. More formally, we write that the acceleration a as $\frac{dv}{dt}$, which is also the same as differentiating $x(t)$ twice, which is written in the form $a = \frac{d^2x}{dt^2}$.

To make this useful, we need two more things. The first is a way to perform this step algebraically. The easiest way of doing this (albeit not the most rigorous) is to use the method of infinitesimals. We define $\frac{dx(t)}{dt} = \lim_{\delta \to 0} \frac{x(t+\delta)-x(t)}{\delta}$. The limit operation means that we gradually reduce δ towards 0; as long as the vector function x is continuous this formula will home in on a particular value. The δ in the denominator becomes small, but so does the numerator at the same rate, so the fraction remains well defined. The formula means that we take the location of the particle (expressed in some suitable coordinate system) at two neighbouring moments in time, find the difference between these two locations and divide it by the difference in time. We then observe how this quantity changes as we bring the two moments in time closer together; it will converge to a particular number, and this number is the differential of x with respect to t taken at that time (the two different moments, of course, by now being so close together that we cannot distinguish one from the other). It is possible for most functions to expand $x(t+\delta) = x(t) + \delta x'(t) + \frac{1}{2}\delta^2 x''(t) + \ldots$, where the $\ldots$ denotes cubic and higher powers of δ, for some functions x', x'' and so on, and these functions are the first derivative, the second derivative, and so on. It is known how to differentiate every mathematical function which is useful in physics.

The second procedure we need is to be able to invert the derivative operator. So we know that if we are given a function $x(t)$ we can find its derivative $\frac{dx}{dt}$; but what happens if we are given the derivative $\frac{dx}{dt}$ and want to find $x(t)$? The process to do this is known as integration. Like differentiation, integration also has a geometrical interpretation: it corresponds to the area under the curve; it also has an algebraic interpretation, namely that it represents the sum of a function evaluated at infinitesimally separated locations:

$$\int_0^a f(x)dx = \lim_{\delta \to 0}[\delta(f(\frac{\delta}{2}) + f(\frac{3\delta}{2}) + f(\frac{5\delta}{2}) + \ldots + f(a - \frac{\delta}{2}))]$$

Here we are considering the function between two points, the first where $x = 0$ and the second where $x = a$. We split the region between those points into various strips of equal size δ, and add up the value of the function at the centre of each strip multiplied by the width δ of the strip. We then consider how this quantity changes as we gradually reduce δ (and correspondingly increase the number of strips); once again this will converge at some particular value which we call the integral of the function over the region between the two points. Pay particular notice to the relationship between integration and summation: it will be important.

However, there are two major difficulties when we try to use integration. The first is that the answer is not unique. That should be obvious from the formal definition of differentiation a few paragraphs ago: if we consider the function $x_1(t) = x(t) + A$ for any constant A, we see that everywhere $\frac{dx_1}{dt} = \frac{dx}{dt}$. This means that whenever we integrate, we can only specify the solution up to this unknown constant A. If we have to integrate twice, we get two arbitrary constants entering our equations.

The second difficulty with integration is that it is much harder than differentiation. While I can (given enough time and patience) differentiate any continuous, well-defined, closed form, and finite function you give me, I can only find an algebraic solution to a relatively small number of integrals. Of course, it is also possible to use computers to solve integrals numerically, but this can only give you specific rather than general solutions (so we can measure the area under the curve $x(t)$ from between $t = t_0$ and $t = t_1$ for specified t_0 and t_1 relatively easily for a computer, up to an in principle arbitrary - but never perfect - precision, we cannot always do so if t_0 and t_1 are not specified but left general. But that is what we want if we are to write down the most general answer to a problem in physics). That doesn't mean that no algebraic solution exists for the integrals - only that we don't know how to find it.

So far, I have been discussing the one dimensional system. What if we extend to three dimensions, which is, of course, what we need to do? The answer to this was provided by Descartes with his coordinate systems and vectors. Instead of writing $x(t)$ as a single function, we write it as a vector, $\vec{\mathbf{x}}(t)$, of three functions; t is just a single coordinate in an unrelated dimension. Calculus doesn't actually care about whether it has one or three functions to worry about: it just proceeds in the same way, albeit with at least three times as much work.

However, when we discuss multiple variables, we need to be a little careful about what happens when they are all interrelated. If we try to change one variable, x, what happens when y and z change at the same time? To solve this, we introduce what is known as the partial derivative. So suppose we have a function $\phi(x, y, z)$, and all of x, y and z are functions of time. This means that if we vary x a little by moving to the next moment in time, y and z would also change a little, and this complicates matters if we want to know how sensitive ϕ is to changes in x alone. For this reason, we introduce the partial derivative, $\left(\frac{\partial \phi}{\partial x}\right)_{y,z}$, which describes what happens if we vary x and (artificially) keep y and z constant.

According to the mechanical principles of Galileo and Descartes, all of physics can be reduced to issues of extension, distance, velocity and acceleration. It has long been known how to treat extension and distance algebraically; Newton (and Leibniz, though Newton was the one who applied this to physics) taught us how to treat velocity and acceleration algebraically. What this means is that every physics problem where mechanical principles apply can be reduced to an equation, and because it involves acceleration, it is a differential equation. A differential equation is simply an equation which includes derivative operators; since the equation for classical physics relates force to acceleration this involves two derivatives for each particle we are considering. This means that to solve this equation, we need to find two constants of integration for each particle under consideration. These constants are not, however, arbitrary in this case. The question we are trying to answer is 'Given that the universe is in a particular state at time 0, what state will the universe be in

at an arbitrary time t?' This means that we know the location and velocity of every particle we are solving for at time 0; these are known as boundary conditions. We can thus (in principle) solve the equations generally, and then find the specific set of constants that give us the right result at time $t = 0$. We have two constants per particle, and two constraints (location and velocity) per particle, so in principle this can be done.

Of course, this leaves two questions: how do we find these equations, and secondly how do we solve them? Neither of these are easy in general. Formulating the equations of motion with a large number of particles is often hard enough; solving them is usually even harder, and in many cases no solutions are known. Even three bodies interacting via Newtonian gravity is insoluble by pen and paper, though it can be done by computer, or by using approximations in certain circumstances. For example, in the case of the solar system, the gravitational interaction between the planets is sufficiently small that to a first approximation we can ignore it. We can then take this first approximation, add a correction to it, and substitute that into our equation of motion, which creates another equation whose solution gives the correction. Once again, this is too difficult to solve, but we can again create a better approximation by removing the less important contributions until we again find an equation we can solve for the correction. This gives us a known first guess and a known initial correction. The sum of these gives a second approximation, which is a better solution than our first guess. But we then need to find the correction to this second approximation, which we can do by the same procedure. Thus by gradually refining our solution we approach the right answer. We will never get there exactly – that will require finding corrections to corrections to corrections an infinite number of times but we can reduce the error to an amount that is smaller than we are likely to measure, and just as importantly we can estimate approximately how large that error is going to be.

This method of finding a series of gradually improving approximate answers is one of the most powerful tools that physicists have. Our first goal is to find an exact solution. If that proves impossible, which is often the case, then we switch to this method. Of course, this method doesn't always work either. We need to be able to split the equation into a part that is important and soluble and a part that is less important that can nonetheless be used to gradually improve our initial solution. We can't always separate the equation in this way. Additionally, sometimes the series of corrections doesn't converge, so the size of the corrections doesn't decrease fast enough as we progress. So if the method of finding successively improvements doesn't work and we can't find an exact solution, then physicists either turn to various models or approximations, where, with a bit of hand-waving, drawings of clouds on a blackboard, and a few too many pints of beer, the physicists manage to persuade themselves that the models capture the essential elements of the physical problem; or we turn to the computer and attempt to solve the problem numerically (for example by duplicating equation (7.1.2) for ever decreasing δ, or by using more sophisticated methods known as Monte-Carlo algorithms). These computer solutions are only at best controlled approximations, and can take up a lot of computer time, but we can at least, if we phrased and programmed the problem correctly (which means if we understood the underlying physics fully), be sure that they will converge towards the right answer and at each step the imprecision of the calculation is known. Models are generally worse than direct computer calculations, principally because it is much harder when using models to systematically and accurately estimate the error of the approximation. However, models, if accurate (which they are frequently not), can give us insight into the understanding of why the world behaves as it does beyond the mathematical calculations.

With regards to finding the laws of motion, Newton left us with a few suggestions. The first three are known as his laws of motion:

(1) A body continues in a state of constant velocity, unless acted upon by an external force.
(2) The acceleration of a body is given by the force divided by its mass.

(3) Every action has an equal and opposite reaction.

The first of these laws is just the principle of inertia, which had already been formulated by the likes of Descartes and Hobbes. The second law needs some qualification. The first question is what is meant by force. A force is something which induces an acceleration. I will discuss specific examples of this in a few paragraphs; here I will just note that, like acceleration, velocity and location, it is represented by a vector: it has both a magnitude and a direction. Then we come to the question of what is meant by mass. In Newtonian mechanics, this is just a property of a body, like extension or colour. We can perhaps (rather naively) think of it as counting how many corpuscles a body is made out of; so, for example, if each corpuscle has the same mass, then (in classical mechanics) the total mass of the body is the mass of each corpuscle times the number of corpuscles that make up the body. If corpuscles come in different types, with each type having a different mass, then it will be the mass of each type of corpuscle multiplied by the number of corpuscles of that type. A dense body would have more corpuscles, and therefore a higher mass; a sparse body of the same size fewer corpuscles and therefore a lower mass. Of course, this just reduces the question of what does the mass of a body mean to what does the mass of a corpuscle mean. Here, the Newtonian ultimately has no answer: it is just a fundamental property of matter that needs to be measured and represents its resistance to acceleration. In Newtonian physics, the mass of a corpuscle cannot change in time, nor the mass of a composite body unless matter is added to or subtracted from that body.

We know that (in classical physics) bodies continue in a state of constant velocity unless acted upon by an external force, and this follows regardless of the masses of the bodies. However, we also know that the same force acting on different bodies can produce different accelerations. If we are to make sense of this, then we must introduce some parameter to describe this different behaviour, and this is what the mass is in Newtonian physics. It measures the amount of force needed to produce a given acceleration. It is found that mass is additive; if we had one body whose mass (as measured by the acceleration under a constant force) was given as one unit[5] of measurement, and another body whose mass was found to be two units, then if we glued these two bodies together we would have a body with mass of three units. It is not immediately obvious that this should be the case, but, rather conveniently, it is with the particular definition of mass given by Newton's second law.

Newton's third law also requires explanation. We need to introduce the concept of momentum, which is mass times velocity. In fact, since mass is constant, the second law can also be re-written in terms of momentum, saying that force is the differential of momentum with respect to time. The third law states that if, in a collision or some over interaction, one body's momentum increases by a certain amount, then the other body must lose the same amount of momentum. This is, of course, means that the total amount of momentum in the universe must remain constant across all time. Whatever is gained in one place, is lost in another by a precisely equal amount.

A further useful concept is that of energy. In Newtonian mechanics, energy comes in various forms. The first of these is potential energy, usually denoted by the symbol V, which is a function of location. The potential energy is related to the force; the force is the differential of the potential energy with respect to location. We usually write $\vec{\mathbf{F}} = -\vec{\nabla}V$, where $\vec{\nabla}V$ is a vector made up of the three derivatives $(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z})$, one derivative for each coordinate representing the three spatial dimensions. We can also introduce the idea of kinetic energy, denoted by K, which is the square of the momentum divided by twice the

[5]This unit could, for example, be kilograms, but we would have to define the kilogram first. In practice what scale we use to measure masses is arbitrary: however the ratio of masses is not arbitrary. We need to fix one particular body as the standard unit of mass, and then define all other masses in relation to it.

mass. The reason we do this is that we find that the second law implies that

$$\frac{\partial}{\partial t}\sum_i (K_i + V_i) = 0. \tag{207.1}$$

This equation needs explaining. The differential with respect to t I have already introduced. $\sum_i$ represents the sum over all particles; $K_i + V_i$ is the kinetic energy plus potential energy of a particle labelled i, and $\sum_i (K_i + V_i)$ measures the total kinetic and potential energy of all particles in the universe. The differential operator measures change, so what this equation is telling us is that the total energy in the universe doesn't change with time. In other words, energy, like momentum, is a conserved quantity.

So far we have just been discussing straight line motion. There is another conservation law related to angular motion. We define angular momentum as $\vec{\mathbf{L}} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}$, where '$\times$' here does not represent multiplication but a particular operator (called the vector or cross product) that acts on two vectors to give another vector, $\vec{\mathbf{r}}$ is a vector denoting the location from the centre of rotation, and $\vec{\mathbf{p}}$ is a vector denoting the momentum. The key thing about this quantity is that the total angular momentum across the universe is also conserved. If you want to stop something rotating, you have to apply a force against that rotation. It is this which causes the earth to continue rotating on its axis. It also explains why rifles are more accurate than muskets. A rifle has various spiral groves in the gun barrel which mean that the bullet, as well as having a great deal of horizontal momentum, is also rotating on its axis so that it carries some angular momentum. If anything tries to alter the course of that bullet, it needs to apply a force to overcome both the horizontal momentum and the angular momentum. A musket ball, however, does not rotate so much when fired, so it just (in effect) has the horizontal momentum. It is thus easier to divert a musket ball than it is a rifle bullet, because it lacks this angular momentum.

Forces can arise from numerous sources. The first of these was well known to Newton's predecessors: physical impact. This creates a force in a direction perpendicular to the surface of impact. The actual details of how this force is applied depend on the materials; the usual vision in classical mechanics is that the transfer of momentum is instantaneous. In practice, of course, this is not the case; each body will compress slightly, and then expand back to its original shape, and it is during this process of compression and expansion that momentum is transferred from one body to another. The precise mechanics of the collision thus depend on the internal atomic bonds between the two bodies. However, the conservation of momentum and energy allows us to compute what happens once the bodies have come into contact and separated. This type of force was not controversial in Newton's day.

The second type of force offers some controversial aspects. This is gravity. Newton proposed that gravity could be explained by supposing that there was an attractive force between every two lumps of matter in the universe, proportional to the product of the masses, and inversely proportional to the square of the distance between them. Thus the force of gravity between two bodies of mass m_1 and m_2 separated by a location vector $\vec{\mathbf{r}}$ which implies a distance r is,

$$\vec{\mathbf{F}} = -G\frac{m_1 m_2 \vec{\mathbf{r}}}{r^3}, \tag{207.2}$$

where G is a constant of nature known as Newton's constant, which basically describes how strong the gravitational interaction is; the minus sign means that this force is attractive. Newton's first breakthrough was to write down this formula. His second breakthrough was to realised that this equation does not only explain why apples fall off trees towards the ground, but was also what kept the planets in orbit around the sun. Newton thus unified terrestrial dynamics and heavenly dynamics in two simple equations (the law of gravity and Newton's second law). This once again stands against Aristotle's physics, which supposes that very different principles apply in the heavens and on the earth.

Newton showed that, from this equation, one could derive Kepler's laws of astronomical motion. The problem of the solar system was thus seemingly solved. Furthermore, we see that the force of Gravity is proportional to the mass. This means that the acceleration is independent of the mass. Thus Newton also explained why when Galileo allegedly dropped two different masses off the tower at Pisa, they hit the ground at the same time. We all emit an attractive force (so a man attracts his wife not only through flowers, but also gravity), but this is too weak to measure for anything except for very massive bodies such as planets (which is why flowers are more effective than gravity). However, a planet will suck everything else in the universe towards it; however this force diminishes rapidly with distance, so that those objects on the surface of the planet will feel a stronger gravitational force than those objects further away. It can be shown for a large spherical body of constant density, the force of gravity for objects outside it is the same as though all the mass were concentrated at its centre. If we are close enough to the surface of a large planet (at much smaller distances than the radius of the planet), the force of gravity will appear to be constant, with the corrections due to the inverse square behaviour too small to easily measure. This again explains what was discovered by Galileo. Weight is defined as the force due to gravity, which is proportional to but not the same as the mass.

Newton's gravity works very well, but it leaves a number of questions unanswered, such as why is the gravitational mass (appearing in the law of gravity) proportional to the inertial mass (appearing in the second law)? And why an inverse square law for gravity? Without this, the solar system would be unstable. More troubling, it seems to go against the traditional mechanistic view that momentum can only be transferred by contact and collisions. The force at a distance aspect of gravity seems to contradict this. Newton himself avoided this problem by appealing to an aether, an invisible fluid permeating all space, which allowed the gravitational and other forces to be explained without recourse to force at a distance. The idea was that the aether possessed certain active powers, such as its individual particles are repulsed by ordinary matter, and grow more dense the further they are away from a planet. This elastic force would propel bodies from regions where the aether was more dense to regions where it was less dense, thus explaining gravity. Other powers of the aether would explain other mysterious phenomena, such as light and magnetism. However, some of Newton's followers rejected this and proposed genuine action at a distance, which seemed to refute the popular forms of mechanism proposed before Newton.

But though it leaves a lot of questions unanswered, Newton's formulation of mechanics works exceptionally well.

There are several main points to note: firstly, Newton began a process of unification of various physical theories; in this case unifying terrestrial dynamics with astronomy. Secondly, Newtonian mechanics depends strongly on the concept of force. Thirdly, Newtonian mechanics is and must be deterministic; if one knows the state of the universe at one instant in time, then if Newton's theory is correct, everything else is fixed across all time.

7.2. From Newton to Einstein

7.2.1. Lagrangian and Hamiltonian Mechanics

There were two reformulations of Newtonian mechanics in the eighteenth and nineteenth centuries. These are mathematically equivalent to Newton's theory - basically just a different way of saying exactly the same thing - but are interesting, partly because they make certain features of the theory clearer, partly because it is easier to perform certain calculations using them, but (in my view) mainly because they provide a useful insight into the bridge between classical and quantum physics. Indeed, Newton's second law is no longer part of our best physical theories. Nor is the concept of *force* as Newton and Galileo understood it. But the Lagrangian and Hamiltonian are still there. Philosophically they are interesting because, while Newton's laws were inspired directly from the mechanical philosophy, these reformulations are not so obviously mechanical. Lagrangian physics was first

proposed as a means of rescuing teleological philosophy from the mechanism that inspired Newton and Galileo; deriving the same physics from a very different metaphysics. It hints that the mechanical philosophy might not be necessary as a basis of classical physics.

The first of these reformulations is known as Lagrangian mechanics, and the definitive formulation was provided by Lagrange in the late eighteenth century[6]. However, the fundamental component, the principle of least action, was first proposed in the early eighteenth century by Leibniz and Maupertuis independently, and most prominently advocated and developed by the noted mathematician Leonhard Euler.

Before discussing these scientists, I need to briefly mention Fermat's work on optics in the 1660s. Fermat discovered (or rediscovered: similar work had been done by Greek and Islamic scientists) that light took the path which travelled from its source to destination in the shortest time. Obviously in a constant medium, this is the path of shortest distance, which is a straight line. But the speed of light is not the same in every medium, so, for example, it is slower in a block of glass than it is in the air, and slower in the air than it is in a vacuum. The reason for this is straight-forward: as the light passes through the medium, it interacts with the various atoms it encounters. It excites an electron in the atom to a higher energy level, and then a short time later the electron decays back to its original energy level and re-emits the photon, the particle that conveys light. Although the new photon can be emitted in any direction, interference effects mean that only those photons travelling the original direction of travel interfere constructively with those photons which missed the atoms. This delay caused by the process of absorption and re-emission causes the effective speed of light in the medium to decrease. So when the light travels from air to glass, the quickest path is no longer a straight line. Since the light travels faster in the air, the quickest route will increase the amount of time the light particle travels in the air. This means that the path of the light bends when it hits the medium, a process known as refraction, and which is well known and well tested experimentally. What this tells us is that optics can be reduced to an optimisation problem, finding the path that minimises the time.

Maupertuis was inspired by the teleological principle that this was the best possible world. This means that the motion of the particles must be ordered in such a way that has the maximum possible fitness for God's purpose. The most natural way of writing this mathematically is to say that a certain function, called the action and denoted by the symbol S, which depends in the paths of the particles, is either maximised or minimised (maximised if this function represents goodness, minimised if it represents badness; in practice we choose the least action principle, although in practice the two are mathematically equivalent: the minimum of S is the maximum of $-S$). Now we know that the path of the particles does not depend on the duration of time they are observed for: we can watch a particle for (say) one second, and observe the path, or we can choose to watch the same particle for two seconds, and observe the path across that time period, and we know that the path travelled over the two second period starts off on the same route as it took for the one second period of observation. Mathematically, if we write S_1 as the action for the first second and S_2 as the action for the second second, then the path travelled minimises both S_1 and S_2. There is obviously nothing special about seconds, so we can see that the action must be minimised for each infinitesimal time slice. The action as a whole can thus be treated as the sum of these infinitesimal time slice actions, and since there is no reason to treat one time slice differently from another, we can write it as the simple sum of the individual actions. So we have $S = S_1 + S_2$, or more generally, if the action on one time slice of duration δt is given by $L\delta t$, we can write this as an integral in the form

$$S = \int dt L(q_i(t), \frac{dq_i}{dt}(t)). \tag{209.1}$$

[6]For an introduction to Lagrangian and Hamiltonian mechanics, see (French and Ebison, 2013; Landau and Lifshitz, 1976; Goldstein et al., 2001)

I have made L into a mathematical function of several variables. In the developed form of the principle of least action, this function is known as the Lagrangian after the man who first wrote down the correct form of it that is equivalent to Newton's original mechanics. L is a function of $q(t)$, which represents the location of the particles, and since there might be more than one particle in the system, I have written this with an index i to represent the locations of each particle separately, so the three numbers q_1, q_2 and q_3 represent the location of the first particle, q_4, q_5 and q_6 the location of the second particle, and so on. These locations are obviously functions of time, since the particles are in motion, and thus I have written them as $q_i(t)$ to show this dependence. Newtonian mechanics, however, needs a second variable to define the state of the particle, the momentum. This is proportional to the velocity of each particle, $\frac{dq_i}{dt}$, and this is also a function of the time, since the particles will experience forces and thus accelerations. Why don't we add the acceleration, and acceleration of the acceleration and so on as well? We could do, but we know from Newton that location and momentum are enough to describe the state of the particle, so we will stick with these. Ultimately (at least in classical physics), the answer is because this is the form of the function that agrees with experiment. Note that this particular formulation assumes that the number of particles remains constant: it cannot cope with particles being created and destroyed.

So what we want to do is find the functions q_i and $\frac{dq_i}{dt}$ which minimise the Lagrangian. Together these are sufficient to describe the location and velocity of the particle at each moment in time and thus solve the chief problem of classical physics, which is to describe and predict the motion of particles. The second thing we need to do is find the form of the function L which correctly describes physics.

The first of these problems was actually solved by Newton. He was interested in the shape taken by a wire which hung between two points, which depends on the tension in the wire, which depends on the angle made by the wire with the vertical at each point and the force of gravity. This was a case of minimising the total potential energy, partly gravitational and partly internal to the wire. This leads to a mathematical problem which is of the same form as the principle of least action, with the potential energy being the sum of the location and gradient of the wire at each point along the wire. As Newton developed the mathematics to solve his own problem, he also unwittingly provided the solution to Euler and Lagrange's problem of minimising the action. The formula is well worth writing out in full.

$$\frac{d}{dt}\frac{\partial L}{\partial \frac{dq_i}{dt}} = \frac{\partial L}{\partial q_i}. \tag{210.1}$$

So what this does is give us a differential equation for each particle describing the motion of each particle. Solve this equation, and we solve the problem. But this also tells us the form of the Lagrangian. Because if we write it as the sum of two terms, the first being the a function of velocity alone, $K = \sum_i \frac{1}{2} m_i \left(\frac{dq_i}{dt}\right)^2$, where m_i is the mass of the ith particle and $\sum_i$ means that we add up the contribution for each particle, and the second a function of location alone, $-V(q_i)$, then the Euler-Lagrange equation written above applied to each particle in turn takes the form,

$$\frac{d}{dt}(m_i \frac{dq_i}{dt}) = -\frac{\partial V}{\partial q_i} \tag{210.2}$$

But the left hand side of this equation is just mass times acceleration. And if we make $V(q_i)$ the potential energy, then the right hand side of this equation is just the force. In other words, the principle of least action with this form of the Lagrangian is equivalent to Newton's second law of motion. If there is no force, then we also get Newton's first law of motion as well (which states that if there is no force, there is no acceleration). Newton's third law also arises from this formulation if we make two further assumptions: firstly, the Lagrangian represents a closed system, so there is nothing outside the Lagrangian which

acts on the particles included in the Lagrangian (the Lagrangian contains every particle in the universe), and secondly that the potential only depends on the distance between the two particles, so it is of the form $V(q_1 - q_4, q_2 - q_5, q_3 - q_6)$ and so on (covering every pair of particles; but for simplicity I will only consider a two particle system). In this case, $\frac{\partial V}{\partial q_1} = -\frac{\partial V}{\partial q_4}$ and so on, which implies the force acting on the first particle is the opposite of the force acting on the second particle. This is just Newton's third law. Of these two conditions, the first is obviously necessary: Newton's law only applies to particles interacting with each other, and not if we just have some vague unspecified force with no particle attached to it. The requirement that the potential only depends on functions of the distance between the particles and not (say) their location in absolute space is also natural because the coordinate system we choose is our own arbitrary choice. This means that physics ought to be independent of where we choose our coordinate origin; since anything except differences of location depends on the origin, this implies that forces should only depend on differences of location.

So the Lagrangian formulation gives us Newton's three laws of motion, and is thus equivalent to Newtonian mechanics. Indeed, since to solve anything in the system we need to apply the Euler-Lagrange equation to give us the Laws of motion, it just adds an extra layer of work. So why use it? Firstly, because it is not always obvious what the equations of motion are, if, like Newton, we just consider forces, or if those forces depend on velocities as well as locations, or for rotating bodies. The Lagrangian formulation provides a way of finding the equations of motion in almost all circumstances. Secondly, it is always nice having an additional way of looking at the problem, because sometimes the new viewing angle can give hints to new ideas. That physics can be expressed in a least action principle is interesting enough in itself. Why should we be able to look at it in this way? Does that give us some clue to the underlying philosophy? The least action principle is not obviously mechanical, which hints that there might be philosophies other than mechanism which can lead us to modern physics.

Another interesting thing about the action is that there are certain transformations of the coordinates which leave the action unchanged. I already mentioned one of them: the shift in the coordinates $\vec{\mathbf{q}} \to \vec{\mathbf{q}} + \vec{\xi}$, where $\vec{\xi}$ represents a constant shift in the location of the coordinates, with all the particles undergoing the same shift. This clearly leaves the action unchanged, with the distance between the particles unaffected, and the velocity of the particles is also left unchanged by a constant shift in location. This is known as a symmetry of the action: a transformation of the coordinates which leaves the action unchanged, and which therefore has no effect on the physics. We can also rotate the coordinates, and perform translations in time, and these are also symmetries of the action. These are all related to our freedom to select any coordinate system we like: rotating the coordinates just means rotating the coordinate system. This is our first introduction of one of the most important and powerful concepts in modern physics. It does not seem like much at the moment (but note that I have just argued that Newton's third law, and consequently the conservation of momentum, is equivalent to saying that the action is symmetric under a shift in location), but I will keep coming back to it, and there are some quite profound conclusions to be drawn. But not yet. There is still some more physics to introduce first.

Lagrangian mechanics was extended further by the Irish mathematician William Hamilton in the early part of the nineteenth century. He wrote down another way of expressing Newtonian mechanics, again not based on Newton's Laws as the most fundamental principle, but equivalent to it. In Newtonian mechanics, and Lagrangian mechanics, the state of the particle was given by its location and its velocity. Hamilton's idea was to use the momentum and location as the two coordinates, but to define the momentum in a particular way based on the Lagrangian. He thus defines a momentum vector π_i as

$$\pi_i = \frac{\partial L}{\partial \frac{\partial q_i}{\partial t}}. \tag{211.1}$$

With the standard definition of the Lagrangian, this is just the standard momentum vector of Newton's theory. However, in more complicated systems, it gives something a little different. The next step is to define a functional, known as the Hamiltonian, as the Lagrangian minus the scalar product of the location and momentum, $H = \sum_i \pi_i \frac{\partial q_i}{\partial t} - L$. Hamilton then showed that the least action theorem, and therefore also Newton's equations could then be reduced to the form

$$\begin{aligned} \frac{\partial \pi_i}{\partial t} &= -\frac{\partial H}{\partial q_i} \\ \frac{\partial q_i}{\partial t} &= \frac{\partial H}{\partial \pi_i} \end{aligned} \tag{212.1}$$

So instead of one second order equation, we have two coupled first order equations.

What is the point of this? The first interesting thing to note is that H represents the total energy of the system; this follows from the definition of kinetic and potential energies. So Newton made location and velocity as the key components of his system, with energy and momentum as something which plopped out of the equations which turned out to be useful. In this formulation, energy and momentum take centre stage, and if they are as fundamental properties as classical calculations seem to imply, then making them central to the formulation is certainly better. The second point is that this formulation leads to some quite interesting mathematical properties, which are beyond the scope of what I want to discuss here. Thirdly, we see that Hamilton's equations give location and momentum nearly symmetric roles. In Newtonian's original formulation, location was the central idea. Momentum was introduced as a by-product of his dynamics. But in this formulation, we see that momentum is something more significant than that. It is in some way analogous to location, and is therefore just as fundamental a part of physics. We can, perhaps, even think of introducing a momentum/ energy space just as we have introduced a space/time space. Equally, both energy and time enter into these equations in a vaguely symmetric way, suggesting that just as location and momentum are analogous, there might also be a relationship between energy and time. Maybe, alongside the usual space/time coordinate system we are familiar with, we should also add in a shadowy momentum/energy as a separate coordinate system mathematically just like it. But most importantly, we have defined momentum in a slightly different way to how Newton originally did. The two methods are equivalent for particles with kinetic energy only dependent on the particle's velocity, and a potential energy independent of the particle's velocity. But there are some systems where life becomes a bit more interesting, as I will shortly discuss. Maybe Newton mislead us about the mathematical form of momentum.

7.2.2. Electricity, Magnetism and Optics

Magnetism and electrostatics have been known from antiquity, though not well understood. The first systematic studies of magnetism were in the medieval and early modern periods, with the best known that done by William Gilbert around the time of Galileo. Gilbert's studies demonstrated the magnetic field of the earth, and also investigated the properties of magnetic iron. Gilbert also clearly distinguished between static electricity and magnetism, and separated these two branches of science. It was then gradually discovered that the electric force decreased with distance, and gradually over the course of the eighteenth century it was discovered that the electric force varied with the inverse square of the distance between the particles, analogous to Newton's theory of gravity (Griffiths, 1998; Jackson, 1998). What is not similar to gravity is that the strength of the force is not proportional to the mass of the particle, but another quality known as electric charge, and that these charges can be either positive or negative meaning that the force can be either attractive or repulsive. The law that the electric force is proportional to the products of the charges of the body, and inversely proportional to the square of the distance is known as Coulomb's law, and was first written down in the late eighteenth century.

Another difference with gravity, however, is that the constant of proportionality is not universal, but depends on the medium between the two different charges.

While electrostatics and gravity may seem to be similar, they do, in reality, arise from very different sources. The inverse square law is perfectly accurate neither in gravity nor electromagnetism, and the corrections to this law are correspondingly different in each case. The analogy between gravity and the electric force is thus accidental, and should not be taken too seriously.

The electric force is in general much stronger than gravity (for example, the electric force between two electrons is vastly larger than the gravitational force). The reason that we feel gravity more than the electric force is gravity is always attractive while the electric force can be attractive or repulsive. So the earth is made of numerous particles, some of positive charge, and others of negative charge, and the net effect of this is that there is almost a complete cancellation between the attractive and repulsive forces; however for gravity all these particles attract other particles towards them.

Like gravity, the electric force seems to act at a distance, and the way this is stated is to introduce a concept of the electric field. The magnitude of the electric field is proportional to the force that would act on a charged particle per unit charge. It emanates from every particle with electric charge, and it is the electric field that reduces over the square of the distance. Another useful concept is the electric flux, which is the amount of electric field passing through an area. So we construct a two dimensional surface, for example the face of a piece of paper. At each point on that surface, one can draw a line parallel to that surface, and another line which is perpendicular to that first line which is also parallel to the surface. So lay your piece of paper flat, draw a straight line, and then another straight line at right angles to it, which intersects the first line at some point in the middle of the paper; you can, of you like, do this more often to make a grid. Now take the piece of paper, and bend and fold it. The two lines will still be there, though no longer straight lines (from the point of view of a Euclidean coordinate system external to the paper) because they will bend along with the paper. At the intersection between the two lines, they give the two directions parallel to the surface of the paper. But we live in a three dimensional space; so there is a third direction, which is the perpendicular to the surface, which comes off the paper at right angles to the two lines drawn on the paper. The direction of the lines parallel to the surface is not unique: you could have drawn them at any angle relative to the sides of the paper. However, there is only one perpendicular line for every point of the surface. This set of perpendicular lines can be used to define the shape of the surface. The electric flux can then be defined as follows: for a given electric field and surface, we take each infinitesimal area on the surface, and measure the component of the electric field in the direction of the perpendicular. We then sum that up over every small area, and that's the electric flux. For a flat surface and constant electric field, the electric flux is just the area of the surface multiplied by the electric field perpendicular to the surface.

Now let us consider a single point of electric charge, and construct a surface around it, a simple sphere with the centre on the electric charge. We now measure the electric flux passing through that sphere, and we find that it is independent of the radius of the sphere. This is straight-forward to show. The electric field drops with the inverse square of the radius, while the surface area of the sphere is proportional to the square of the radius. Multiply them together, and the dependence on the radius for each factor cancels out, and you just get a constant, and this constant is proportional to the electric charge of our point source. It is conventional to call the charge Q and, in a vacuum, the constant of proportionality as $1/\epsilon_0$. In mathematical terms we then have,

$$\int_S \vec{\mathbf{E}} \cdot \mathbf{d}\vec{\mathbf{S}} = \frac{Q}{\epsilon_0} \tag{213.1}$$

The left hand side of this is just the definition of the electric flux expressed in the notation of calculus. $\mathbf{d}\vec{\mathbf{S}}$ is a vector in the direction of the perpendicular to the surface with the

magnitude of an infinitesimal unit of area. $\vec{\mathbf{E}}$ is the electric field, and $\cdot$ just means that we take the component of $\vec{\mathbf{E}}$ that is in the direction of $\mathbf{d}\vec{\mathbf{S}}$ and multiply it by the magnitude of $\mathbf{d}\vec{\mathbf{S}}$, which is the small infinitesimal area. $\int_S$ means that we sum up the results from each area across the whole of the sphere. The right hand side is just the electric charge divided by ϵ_0, our constant of proportionality.

Now it is time to re-introduce the mathematical construct called $\vec{\nabla}$. This is just a vector made up of the differential operator in each direction. So, in Cartesian coordinates, $\vec{\nabla} \equiv (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$. One can combine $\vec{\nabla}$ with a vector to give a scalar property, an operation known as taking the divergence and denoted as $\vec{\nabla} \cdot \vec{\mathbf{E}}$. We can write the three components of the electric field as $\vec{\mathbf{E}} = (E_x, E_y E_z)$, and then we define

$$\vec{\nabla} \cdot \vec{\mathbf{E}} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}. \tag{214.1}$$

There is a mathematical theorem which states that

$$\int_V \vec{\nabla} \cdot \vec{\mathbf{E}}\, \mathbf{d}\vec{\mathbf{S}} \cdot \mathbf{d}\vec{\mathbf{r}} = \int_S \vec{\mathbf{E}} \cdot \mathbf{d}\vec{\mathbf{S}}. \tag{214.2}$$

Here $\mathbf{d}\vec{\mathbf{r}}$ represents an infinitesimal line segment along the perpendicular to the surface. Thus $\mathbf{d}\vec{\mathbf{S}} \cdot \mathbf{d}\vec{\mathbf{r}} \equiv dV \equiv dxdydz$ means an infinitesimal element of volume, and $\int_V$ means that we sum up the value of $\vec{\nabla} \cdot \vec{\mathbf{E}}$ for each infinitesimal unit of volume.

Let's now look at the divergence of the electric field from a static point source with charge Q,

$$\vec{\mathbf{E}} = \frac{Q\vec{\mathbf{r}}}{4\pi\epsilon_0 r^3}. \tag{214.3}$$

where $\vec{\mathbf{r}} = (x, y, z)$ and $r = \sqrt{x^2 + y^2 + z^2}$ (which is just a representation of Pythagoras' theorem). $\vec{\mathbf{E}}$ is infinite at $r = 0$, precisely where the charge is, so lets leave that to one side for the moment. Everywhere else in space it is easy to compute the divergence of the electric field, and we get $\vec{\nabla} \cdot \vec{\mathbf{E}} = 0$. But that can't be right. We have already shown that for a spherical volume $\int_V \vec{\nabla} \cdot \vec{\mathbf{E}} dV = Q/\epsilon_0$, regardless of the value of the radius r. What this means is that we need to consider that point at $r = 0$ a little more closely. I will introduce the quantity known as the Dirac δ-function, $\delta^3(\vec{\mathbf{r}})$, which is basically infinite at $\vec{\mathbf{r}} = 0$ and zero everywhere else. More precisely, if we can construct an infinitesimal volume around $\vec{\mathbf{r}} = 0$, then $\delta^3(\vec{\mathbf{r}})$ is inversely proportional to that volume within the volume, and zero everywhere else. This means that $\int_V dV \delta^3(\vec{\mathbf{r}}) = 1$ if the integration volume contains the point at $r = 0$ and 0 if it doesn't. We thus can write,

$$\vec{\nabla} \cdot \vec{\mathbf{E}} = \frac{Q}{\epsilon_0} \delta^3(\vec{\mathbf{r}}). \tag{214.4}$$

Now, we introduced it as a point-like charge, and that is why this delta function is appearing in the equation, because the δ-function mathematically represents a point like object. In practice, this might make us a little nervous: maybe the fundamental particles aren't point like, but (as Descartes suggested) extended over a small area. That's fine. We can introduce a charge density, conventionally labelled $\rho(\vec{\mathbf{r}})$, which represents the charge contained in a given area. In that case, it is only a short step to get

$$\vec{\nabla} \cdot \vec{\mathbf{E}} = \frac{\rho(\vec{\mathbf{r}})}{\epsilon_0}, \tag{214.5}$$

and this is known as Gauss' law after the brilliant mathematician who first wrote a form of it down. Incidentally, this equation means that equation (213.1) is not just satisfied for spheres with point-like charges at the centre: it is true for any volume of any shape whatsoever with any pattern of electric charges, where Q now represents the total charge contained in the volume.

Now, let us turn our attention over to magnetism, and the nineteenth century scientist Michael Faraday, who was one of the first people to systematically investigate the relationship between electricity and magnetism. Electric charges in motion can be quite useful; indeed, almost everything in our modern world is dependent on electricity. Faraday noted that electric charges in motion interacted with a magnetic field. Indeed, he discovered that a magnetic field imparted a force to a moving, but not static, electric charge. The strange thing is that this force is at right angles to both the particle's velocity and the direction of the magnetic field (which can be measured by placing magnetic iron shavings in the presence of a magnet: they align themselves with the magnetic field). Towards the end of the nineteenth century, due to Thompson, Heaviside and Lorentz, the correct form for the the force provided by a magnetic field $\vec{\mathbf{B}}$ on a charge Q travelling with velocity v was discovered. It is

$$\vec{\mathbf{F}} = Q\vec{\mathbf{v}} \times \vec{\mathbf{B}}. \tag{215.1}$$

Like the electric field, the magnetic field is something which permeates all space, has its origins in moving electric charges, and which generates a magnetic force. For the moment, we can use equation (215.1) as the definition of the magnetic field in terms of the force experienced by a moving particle. The vector product, $\times$, takes in two vectors as input, and spits out another vector as output. That vector is at right angles to both $\vec{\mathbf{v}}$ and $\vec{\mathbf{B}}$, as observed for the magnetic force, and has the magnitude of $\vec{\mathbf{v}}$ multiplied by the component of $\vec{\mathbf{B}}$ in the direction that is at right angles to both $\vec{\mathbf{v}}$ and $\vec{\mathbf{F}}$. In practice this means that if $\vec{\mathbf{F}} = (F_x, F_y, F_z)$, $\vec{\mathbf{v}} = (v_x, v_y, v_z)$ and $\vec{\mathbf{B}} = (B_x, B_y, B_z)$ then $F_x = Q(v_y B_z - v_z B_y)$, and similar for the other components. Similarly, we can define a differential form of the vector product, called the curl and (when acting on a vector $\vec{\mathbf{A}}$) denoted by $\vec{\nabla} \times \vec{\mathbf{A}}$.

Faraday in his experiments on induction noted that a magnetic field can interact and generate an electric current. This is the principle used in transformers, but of particular interest to us is his observation that a changing magnetic field can induce an electric current, i.e. set up an electric potential. In particular, Faraday noted that the rate of change of magnetic flux was equal to the electric potential generated. With potential being electric field times distance, this leads to the integral equation

$$\int_C \vec{\mathbf{E}} \cdot \vec{\mathbf{dl}} = \frac{\partial}{\partial t} \int_A \vec{\mathbf{B}} \cdot \vec{\mathbf{dS}}, \tag{215.2}$$

where $\vec{\mathbf{dl}}$ is an infinitesimal element of the line C that bounds the area A through which we measure the flux $\vec{\mathbf{B}} \cdot \vec{\mathbf{d^2S}}$. This is most commonly quoted in a different form using another mathematical identity,

$$\vec{\nabla} \times \vec{\mathbf{E}} = \frac{\partial}{\partial t} \vec{\mathbf{B}}. \tag{215.3}$$

Electric currents are most commonly associated with the name of the French scientist Andre-Marie Ampére, and the familiar measurement of current, the Amp, is named after him. An electric current is the flow of electrically charged particles; we now know that in the applications we use these particles are electrons. The Amp is a measure of the amount of charge passing through a surface in a given unit of time, and this current is given the symbol $\vec{\mathbf{J}}$. It is, of course, a vector, since like any flow of particles it has both a direction and magnitude. Now just as a magnetic field can induce an electric current, so can an electric current induce a magnetic field. So if we have electrons travelling in a thin and straight wire, there will be a circular magnetic field surrounding that wire. Ampére discovered the law governing the strength of the magnetic field compared to the strength of a unchanging electric current, which was again given in integral form

$$\int_C \vec{\mathbf{B}} \cdot \vec{\mathbf{dl}} = \mu_0 \int_A \vec{\mathbf{J}} \cdot \vec{\mathbf{dS}} \tag{215.4}$$

μ_0 is another constant of proportionality, otherwise the symbols mean the same thing as in Faraday's law. James Clerk Maxwell, however, realized that this equation was incomplete, and that it broke down if the electric current, and in particular the electric field, was not constant in time. In the alternative form, the correct equation is

$$\vec{\nabla} \times \vec{\mathbf{B}} = \mu_0 \vec{\mathbf{J}} + \mu_0 \epsilon_0 \frac{\partial}{\partial t} \vec{\mathbf{E}}. \tag{216.1}$$

There is one equation missing. We have electric charges, but not the equivalent for magnetism. This hypothetical object is known as the magnetic monopole, and it has never been observed. In that case, we have an equivalent to Gauss' law but since there are no magnetic monopoles, the right hand side of the equation is zero. Thus we have,

$$\vec{\nabla} \cdot \vec{\mathbf{B}} = 0. \tag{216.2}$$

We have introduced five equations which are so important, concise, and beautiful that they are worth writing out again,

$$\vec{\nabla} \cdot \vec{\mathbf{E}} = \frac{\rho}{\epsilon_0} \qquad \vec{\nabla} \cdot \vec{\mathbf{B}} = 0$$

$$\vec{\nabla} \times \vec{\mathbf{B}} - \mu_0 \epsilon_0 \frac{\partial}{\partial t} \vec{\mathbf{E}} = \mu_0 \vec{\mathbf{J}} \qquad \vec{\nabla} \times \vec{\mathbf{E}} - \frac{\partial}{\partial t} \vec{\mathbf{B}} = 0 \tag{216.3}$$

$$\vec{\mathbf{F}} = q(\vec{\mathbf{E}} + \vec{\mathbf{v}} \times \vec{\mathbf{B}}) \tag{216.4}$$

The top four equations were first written down together by Maxwell, and are known as Maxwell's laws of electromagnetism. The bottom equation was first written down by Lorentz and is known as the Lorentz force law. Everything, and I do mean everything, in the classical theory of electricity and magnetism is contained within these five equations. The top two equations on the left describe how electric and magnetic fields are generated by static and moving electric charges. The two equations on the right state that there is no similar particle which generates a magnetic field when at rest and electric field when in motion. These four equations together also describe how electric and magnetic fields vary across space. Finally, the last equation describes how the fields interact with and affect the motion of matter. These laws as I have written them are valid in a vacuum. Where there is not a vacuum, the laws are slightly modified because the constants ϵ_0 and μ_0 take a different value which depends on the medium.

There were many great minds who brought classical physics to its culmination; Newton who started it, and Einstein who brought it to a conclusion are rightly well known. James Clerk Maxwell deserves a place alongside them; he was, in my view, the pre-eminent genius of the nineteenth century (obviously biologists would nominate Darwin; and there were many great mathematicians and musicians as well; but, from my admittedly biased perspective, I would place Maxwell above them all[7]). Alongside his work in electromagnetism and optics, Maxwell's greatest achievements were in the fields of thermodynamics and statistical physics (which he had a hand in founding).

These five equations were each drawn from experiment, and two and a half of them were developed by people before Maxwell (though it was Maxwell who expressed them in the differential form I have given above), and the fifth, though Maxwell contributed to it, by someone after him. So why are they so great? From the perspective of modern physics, it is noteworthy that they can be expressed in terms of a Lagrangian for the electromagnetic potentials, and this Maxwell Lagrangian (with the concepts you need to construct it) is just about the only thing from classical physics that survives into the standard model of particle physics. Maxwell's work, in a slightly modified (but in classical mechanics mathematically

[7]Almost everything second and third year physics undergraduates learn is in a subject to which Maxwell either directly contributed, or whose development was inspired in part by his work. They are hit over the head with his brilliance time and again. Thus all physics undergraduates fall in love with the guy, including me.

equivalent) form, is still part of our most fundamental theory today. This isn't true for any physics that Newton did (and almost every other pre-twentieth century scientist, with the exception of Lagrange and those such as Boltzmann who worked with Maxwell on statistical physics), and so established is it that I cannot ever see it being overthrown. But that's getting ahead of myself. Why are they so great in classical physics? The first thing is that they unify the theories of electricity and magnetism, which in ancient times were seen as separate phenomena, and expresses them in simple mathematical form. But the second thing, and quite unexpected, is that they also contain all of optics as well. There will be more to come later, but I will talk about optics first.

The study of light has, of course, entranced scientists since the beginning of methodological investigations. Several basic principles were discovered:

(1) Light travels in a straight line when travelling through a medium.
(2) When it encounters an object, some of the light is absorbed if the medium is opaque or translucent, while some is reflected off that surface and some refracted.
   - Reflected light is done in such a way that the angle of incidence is equal to the angle of reflection.
   - Refracted light continues through the medium, but at a different angle. If the light is travelling in a direction $\vec{\mathbf{k}}$ when it hits the surface of the medium, then the components of $\vec{\mathbf{k}}$ parallel to the surface are reduced by a constant factor which depends on the type of substance and the colour of the light. This is known as Snell's law.
(3) When passing through a small slit, the light comes out in bands of high and low intensity. This is known as diffraction. A different pattern of bands emerges if light travels through multiple slits.

Diffraction is a property familiar to studies of waves. A wave (for example in water) is something which transmits energy and momentum over long distances without any particles moving more than a small amount. In water waves, one particle transfers energy to the particle next to it, and so on as the energy gradually transverses the medium. The shape of the wave at a frozen moment in time is (roughly) a periodic function of location, with successive peaks and troughs. We can imagine a wave as a distortion of the flat surface. Water is taken out of one part of the surface, to give a trough, and placed where there is a peak. However, in practice the individual water particles in the waves only move up and down as the motion of the wave moves across the water. We take the average height of the water as the zero point, and the amplitude of the wave, which can be positive or negative, represents the maximum distance of the water above or below that average. However the energy carried by the wave is not proportional to the amplitude, but rather the square of the amplitude, known as the intensity.

This has an important effect when two or more waves cross each other from different directions. Because when the waves encounter each other, the total height of the new pattern is the sum of the amplitudes. So let us say that there are two waves with amplitudes $A_1(\vec{\mathbf{x}}, t)$ and $A_2(\vec{\mathbf{x}}, t)$, which are, of course, functions of location and time, and intensities $I_1(\vec{\mathbf{x}}, t) = A_1(\vec{\mathbf{x}}, t)^2$ and $I_2(\vec{\mathbf{x}}, t) = A_2(\vec{\mathbf{x}}, t)^2$. The total amplitude is $A_T(\vec{\mathbf{x}}, t) = A_1(\vec{\mathbf{x}}, t) + A_2(\vec{\mathbf{x}}, t)$ which means that the total intensity is $I_T(\vec{\mathbf{x}}, t) = I_1(\vec{\mathbf{x}}, t) + I_2(\vec{\mathbf{x}}, t) + 2A_1(\vec{\mathbf{x}}, t)A_2(\vec{\mathbf{x}}, t)$. This means that the total intensity is not the sum of the intensities of the two individual waves, but something different. The total energy at a given point is not the sum of the energy of the two individual waves. This is the most important difference between waves and particles. The total energy of the sum of a system of particles is the sum of each individual particle energy. With waves this is not the case.

In particular, with waves we can have interference effects. If the two different amplitudes are both positive or both negative, then the total energy will be greater than the sum of the two individual waves. If the two different amplitudes have opposite signs, then the total energy will be less than the sum of the two individual energies. It is even possible that there

can be zero energy at certain points when two waves, which both individually carry energy, cross paths.

Each wave travels in a certain direction at a certain speed. We thus need a vector to denote the direction, which is called $\vec{\mathbf{k}}$, and we also introduce the frequency ω which tells us how fast (in time) it oscillates between peak and trough. The location and time dependence of a peak of a constant-velocity wave can therefore be written as $A(\vec{\mathbf{x}}\cdot\vec{\mathbf{k}}-\omega t)$, so amplitude is a function of time and space together in this particular combination. This means that every wave peak with constant velocity satisfies the equation

$$\vec{\nabla}^2 A - \frac{1}{v^2}\frac{\partial^2 A}{\partial t^2} = 0, \tag{218.1}$$

where the speed of the wave $v^2 = \omega^2/(\vec{\mathbf{k}}\cdot\vec{\mathbf{k}})$. The wave-length of the light is the distance between successive peaks, and this is proportional to $1/\sqrt{\vec{\mathbf{k}}\cdot\vec{\mathbf{k}}}$ (the constant of proportionality is in this case undefined, since I haven't defined precisely what the function A is).

Diffraction effects occur when we have two different waves coming from two different directions interfering with each other. Because the waves come from different directions, the peaks and troughs of each wave cannot be in perfect agreement, and there will be places where there is constructive interference, and other places where there is destructive interference. One well known case where we have interference is the two slit setup, shown in figure 1. In this case, we can treat each hole in the slit as a separate light source, which emits light in every direction. Light travels in a straight line from the source to the hole, then in a straight line from each slit to the observer. The light seen by the observer is thus the contribution from both of these paths: we add together the amplitudes. But because the two paths taken by the light to reach the observer have different lengths, each beam of light will be at different points in its oscillation. This means that as we move the location of the observer, we will encounter regimes where the two beams interfere constructively and regimes where they interfere destructively. Hence we see successive light and dark bands. We also get effects from a single slit. In the above discussion, I assumed that the holes were just points, to keep things simple. In practice they aren't: each slit will have a small but non-zero width. What we then do is treat each point between the start and end of the slit as an individual point source of light. Light travelling from the top of the slit will travel a different path to light travelling from the bottom of the slit, and every point between them a slightly different path. The path lengths are slightly different, and this is enough for the observer to see interference between the different beams of light. This means that even from a single slit, you get bands of constructive and destructive interference. Of course, in reality, we have to take into account both of these effects when we have a two (or more) slit system; but this can be accounted for easily if we go into full mathematical detail (which I don't want to do here).

Of course, this whole discussion begs an important question: why can we treat the slits as separate point sources of light, emitting light in every direction, rather than just the direction the light originally came in? This hypothesis is known as Huygens' principle, but it was originally induced from experiment rather than derived from more fundamental issues. The answer to this question is both surprising and immensely important, but you are going to have to wait several more chapters before I give it. But keep this question in mind, because I will come back to it. (Another important and related question is 'Why does light travel in straight lines?' You should be asking this as well. But again, it is going to be some time before I provide the answer.)

Diffraction and interference effects in general are the crucial signal of wave-like behaviour: if you see diffraction, then you have a wave, and it occurs for all waves. This is again going to be an important issue which I will take up again later.

So that is a quick summary of what was understood about optics at the time of Maxwell. Light was confirmed as a wave, with the different colours referring to different wavelengths

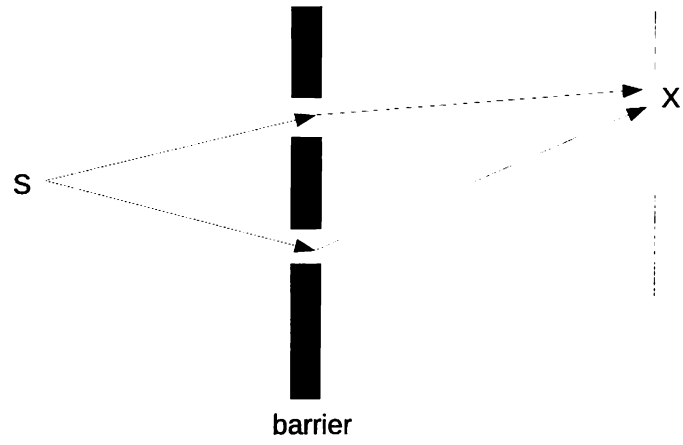


FIGURE 1. A single light source at location S emits light which hits an impenetrable barrier with two holes in it at A and B. The light, whose path is denoted by the arrows, then travels onwards to the point of observation X. The distance travelled by each beam of light is different depending on which path it takes, which means that the two wavefronts might be out of phase at X. We will therefore get either get either constructive or destructive interference depending on the precise distances involved and the wavelength of the light.

of light. The laws of refraction, reflection and diffraction were known and the mathematical description of the various effects well-established. It was also known that light was polarised (which I haven't discussed), i.e. different light waves could oscillate in different directions: for example the oscillation could be up and down, or left and right, and these could be distinguished separately. The speed of light in a vacuum and various different mediums had been measured to a reasonably good precision.

One important question which was not answered was what was the medium through which the light waves oscillated? Water waves had water, sound waves had air, so what was the equivalent for light? After all, it was thought that everything could be ultimately reduced to particles in motion, and this certainly worked for the other known types of wave. The corresponding medium for light waves was called the aether, and was assumed to be an invisible fluid that filled the universe. Of course, giving it a name doesn't explain what it is, and this was a question which worried people.

So what does this have to do with Maxwell and his equations for electromagnetism? Maxwell noticed that, in a vacuum, so without any charges or current, he could combine his equations in such a way that they gave a wave equation for the electric and magnetic fields. These waves would be polarised, and their speed, which could be computed from the constants ϵ_0 and μ_0 measured from electromagnetic effects without any reference to optics, corresponded perfectly with the previously measured speed of light, c. Furthermore, his equations could be used to explain why the speed of light changed in a medium, and to derive the laws of reflection and refraction, giving not only the angles but also the amplitudes: how much of the wave was reflected, and how much refracted. In other words, Maxwell's equations predict, and contain with them, that there would be a wave in the electric and magnetic field whose properties exactly matched what was already known about light. It is not hard to conclude that if Maxwell's equations predict that something should be there, and we see something there with precisely the same properties, then we can equate the two. Furthermore, they predicted that visible light was just one small set of wavelengths associated with the electromagnetic spectrum, and there were other forms of electromagnetic

radiation (which we now call radio waves, infra-red, ultra-violet, x-rays, and so on, although in reality they are just a continuous spectrum). In other words, Maxwell's equations plus the Lorentz force law not only describe everything to do with the electrical and magnetic forces, but also all of optics. These three different areas of physics were unified together in one stroke.

The idea of a unification of physics, that everything could be reduced to one principle, is an appealing one to physicists. Almost all of modern physics has been moving further towards that goal. Newton unified the terrestrial and celestial theories of dynamics and gravity; scientists of the nineteenth century (with Maxwell himself playing a prominent role) unified thermodynamics (the study of heat), with Newton's theory of motion. Einstein unified gravity and motion. Today, of course, we are reduced to just two theories, the standard model of particle physics, and the standard model of cosmology, which between them seem to describe everything. The hope is that these two can be combined into a single theory. The unification by Maxwell (with his predecessors) of electricity, magnetism and optics, and, with the Lorentz force law its unification with Newton's mechanics, was an important part of this trend, and is one of the reasons why at the close of the nineteenth century physicists thought that they were closing rapidly in on the final answer. Little did they know the revolutions that were around the corner.

There is one more important aspect of electromagnetism, which is the magnetic potential. Similarly to the case of gravity, there is an electric potential, which I will denote ϕ, whose differential gives the electric field and which represents the energy of the field. Is there something similar for magnetism? The difficulty is that the magnetic force affects matter in a different way to electricity and gravity. We can't just have a simple $\vec{\mathbf{B}} = \vec{\nabla}\phi_B$ for some function ϕ_B as we have $\vec{\mathbf{E}} = \vec{\nabla}\phi$, if nothing else because in view of Maxwell's law $\vec{\nabla}\cdot\vec{\mathbf{B}} = 0$ the only allowed solution would be $\phi_B = 0$. But there is another option consistent with Maxwell's laws, which is to introduce a three component vector potential $\vec{\mathbf{A}}$ and write $\vec{\mathbf{B}} = \vec{\nabla}\times\vec{\mathbf{A}}$. This potential is a three vector, similar to location and momentum, while the electric potential ϕ is just a single number, like energy or time. So all of the six numbers contained within $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ can be expressed in terms of the four numbers $\vec{\mathbf{A}}$ and ϕ. What we observe are the accelerations on the particles, from which we can directly deduce the electric and magnetic fields; in other words we can measure $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$. We cannot measure ϕ or $\vec{\mathbf{A}}$: they do not correspond to any observable quantity, neither in classical mechanics nor even in quantum mechanics. So are they just mathematical fictions?

It gets even worse: $\vec{\mathbf{A}}$ is not unique (and this will turn out to be crucially important): it has something called a gauge symmetry. If we perform the transformation

$$\vec{\mathbf{A}} \to \vec{\mathbf{A}} + \vec{\nabla}\alpha \tag{220.1}$$

for any function of space α the magnetic field is left unchanged. But the magnetic field is what we observe. There is therefore a redundancy in the description of $\vec{\mathbf{A}}$; numerous different values of $\vec{\mathbf{A}}$ describe exactly the same physics.

So are these potentials just mathematical fictions? We should be slow to dismiss them so quickly. One of the features of physics is that if something emerges in the equations, it almost always turns out to have a key role to play in real life. As it turns out, they are not fictions. There is an experiment we can perform which shows that the potentials do play a fundamental role. However since this experiment relates to quantum mechanics rather than classical physics, I will delay my discussion of it.

Next, I need to discuss how magnetic fields can be expressed in the context of Lagrangian mechanics. The Lagrangian for gravitational interactions was straight forward, the kinetic energy minus the potential energy. For the electrostatic potential, we can use something similar, but that won't work for the magnetic potential, because the force depends on the velocity of the particle. Instead, to get it to work, we need to modify the kinetic term, and add the magnetic potential to the momentum. The correct Lagrangian which generates the

Lorentz force equation for a particle of charge Q is

$$\mathcal{L} = \frac{1}{2m}(m\frac{d\vec{\mathbf{q}}}{dt} - Q\vec{\mathbf{A}})^2 - Q\phi - m\phi_G - \int d^3x \frac{\epsilon_0}{2}(\vec{\mathbf{E}}\cdot\vec{\mathbf{E}} - c^2\vec{\mathbf{B}}\cdot\vec{\mathbf{B}}), \tag{221.1}$$

where ϕ_G is the gravitational potential. Minimise the action with respect to the particle coordinates $\vec{\mathbf{q}}$ and we get the Lorentz Force Law and Newton's law of gravity; minimise with respect to the electromagnetic potential ϕ and $\vec{\mathbf{A}}$ and we get Maxwell's equations. If we look at Hamiltonian's equations, then the canonical momentum is not mass times velocity, but rather mass times velocity minus the charge times the magnetic potential. This is interesting, because it establishes that $\vec{\mathbf{A}}$ is in some way analogous to a momentum, while the electric potential ϕ is obviously a type of energy. But this also raises the question why this is the case. This is ultimately an experimental result; in classical physics it is just a brute fact with no real explanation. Once again the explanation arises quite beautifully when we consider quantum physics, but that is a discussion for later.

Finally, since I started with a discussion of magnets, I ought to make a brief note to finish up that thread. Magnets naturally generate a magnetic field, but Maxwell's equations discuss the movement of electric charges. However in naturally occurring magnets there is no electric current. So what is happening here, and can we merge it into our current picture? Now each subatomic particle has an additional property known as spin. Atoms are built up of these particles, and thus also contain spin. A Newtonian physicist can reason as follows: spin can be explained by thinking of these particles as rotating discs or spheres with electric charge evenly distributed around them. The rotation (or *spin*) of the atoms thus causes the electric charges on the edge of the atoms to move in a circular orbit, and thus generate a magnetic field. Each atom therefore generates a small magnetic field. Now these magnetic fields can either point in the same direction throughout the material, or there might be regions of local coherence, but each region (or *domain*) pointing in a different direction. In the first case, the tiny magnetic fields from each individual atom add together, and in total you get a large magnetic field, with the result being a magnet. If the individual atoms are spinning in different directions, then you get a cancellation between the magnetic fields from different atoms, and the lump of iron (for example) would not be magnetic. Thus the classical physicist could explain magnets, as long as he ignored the question of why the atoms never slowed down and eventually stopped their rotations under the influences of external electromagnetic forces.

This classical explanation is only partially correct. It is true that each individual atom (for certain materials) is like a small magnet, and everything discussing what happens at distances greater than the atomic scale in that explanation is broadly correct (albeit missing a lot of interesting detail). Spin, however, has nothing to do with the rotation of electric charges or the rotation of fundamental particles at all. It arises from an entirely different principle, and the reasons behind it again lie deep in quantum mechanics (and in this case both quantum mechanics and relativity together). The name 'spin' is very unfortunate, because it gives a misleading impression of rotation. Unfortunately, it has been in use for too long and we are now stuck with it.

Maxwell's equations are important, not only for the physics that they describe, but because they provide a bridge between Newton's physics and the twentieth century physics; electromagnetism and optics provided clues for the development of quantum mechanics. They were the last great work of the Newtonian era in physics. But they also add to the many unanswered questions of classical physics. Why Huygens' principle? What is spin? Why does the magnetic field arise as it does? Why don't magnets degrade in time as the atoms stop rotating (and although spin arises from a different source, there is still something important in this question)? Thus electromagnetism also might make us think that classical physics isn't the real answer to the problem. But the influence of Maxwell's equations was greater still as an influence of Einstein and his development of the theory of relativity. Because Maxwell's equations are fully consistent with relativity; and indeed relativity helps

us to understand the links between electricity and magnetism far better than in Newtonian physics.

One other effect of Maxwell's equations, however, is more philosophical. Before Maxwell, a naive form of mechanism had been prevalent: all that mattered was particles in motion, and these affected each other through contact alone. Basically, two particles collide with each other, and change how each other travel. Newton's gravity, with its action at a distance, is obviously something slightly problematic here; instead we have a gravitational force field. Maxwell, following up a suggestion from Faraday, showed that electricity, magnetism and even optics were also all associated with a field, which can be expressed either as a force field spread out over all space or as a potential spread out over space of which the force is the derivative. The field also carries energy, and momentum.

What this means is that the contact picture of mechanism is wrong. Instead, in classical physics, we have two separate entities: particles, which are localised at a single point and represent matter, and fields which are de-localised and carry the interactions between matter. The concepts of field is something not immediately obvious to us, and it took time for it to be accepted. Between them, Newton and Maxwell described all the then known forces in terms of fields. Since that time, we have added two more forces: the weak and strong nuclear forces, and again these are described in terms of a field description. But equally, the field picture removes the action at a distance that troubled philosophers about Newton's work. It is not that one electric charge affects another electric charge a distance away from it, but one electric charge generates an electric field, which smoothly evolves as it goes out into space, and encounters another charge at that location. The interaction between the electric charges and the fields is entirely localised; the change in the electric field is again a local phenomena, since it is smooth and described only by differential operators. There is no interaction at a distance; only contact between particles and fields, and also between one part of a field and the part right next to it.

This means that the naive mechanism which inspired Galileo and sparked the scientific revolution is wrong. But it does not imply that we need to abandon mechanism altogether. The major points of mechanism – indestructible matter which follows deterministic law – remains in Maxwell's physics. But the difference is that there is an intermediary: localised particles of matter are not all there is, and the fields should be assigned a real existence alongside it as a different type of being. But it it is still a mechanistic world view: just a different type of mechanism. In classical physics, particles are localised and discrete; fields are extended (de-localised) and continuous[8]. Particles have properties which can be directly observed; fields only reveal themselves via their influence on particles. Particles interact with fields only and never other particles; fields are generated by both particles and changes in other fields[9]. Of course, it might be thought that the fields are just a mathematical convenience, but this is not the way to think about any physics. If it is in the equations, almost certainly it corresponds to some reality.

Maxwell's physics represented one other change in the approach to physics. Before Maxwell, most of the physical rules concerning electromagnetism and optics had been developed by noticing patterns in experiments. Maxwell (and subsequently Hertz who tidied up much of Maxwell's work) had a similar role here as Newton had in astronomy. Kepler had devised his Laws by pouring over Tycho's detailed observations searching for patterns, trying one thing after another until he found something that worked. Newton found a basic

[8]Discrete means that you can have one particle, or two, or three, but not half a particle – if you split a particle into two, you wind up with two particles, not two halves. Continuous means that the field strength or potential or amplitude can take any real number.

[9]For example, an electric field is generated from a changing magnetic field; although this view is changed once we introduce relativity and consider the potential, rather than the field, as the fundamental object. However, the idea that force carrying fields can interact with other force carrying fields is reasserted when we introduce the nuclear forces. But by that point, we are well beyond classical physics anyway. And for any physicists reading this, I do know about the Higgs Boson – but this is a discussion about classical physics.

mathematical principle that explained Kepler's laws. Gilbert, Coulomb, Ampere, Faraday, Snell and the rest of the early pioneers in electromagnetism and optics had identified patterns from experiment (perhaps Huygens with his explanation of diffraction was the exception); Maxwell explained those patterns in terms of a formal and abstract mathematical structure. But the most important difference was that the earlier studies focused on objects we could, as it seemed, directly observe: copper wires, electrical charges, and the like. Maxwell (following Faraday) instead stated that something we could not, as it seemed, directly observe was equally real and equally important, namely electric and magnetic fields. The fields are ultimately just a mathematical structure: we don't observe them, but only their effects, and yet in Maxwell's theory of electromagnetism they are the central players. Mathematical formalism, rather than observation, is used to identify one of the most important parts of the physical world. Maxwell thus began the process of turning theoretical physics from a study where the sensible objects were paramount to one where the underlying abstract mathematical structure was of primary importance; and the application to sensible objects merely a consequence of the underlying structure. Maxwell marked the beginning of the shift from empiricism to a formalism, which became increasingly important as the twentieth century progressed.

I would like to think about how this affects Aristotle's metaphysics. It might be thought to be one further blow against Aristotle. Aristotle said that there was just matter and form; but now (even if we grant that much) there now seems to be matter, form and fields. However, this I think, takes things too far. We think of matter in terms of localised particles, but there is no reason why we need to. An Aristotelian who adopted a realist location about fields would have to state that they were just a different type of matter, presumably with its own type of form along with it. Or they could just deny that fields have a real existence. But we should not use our modern nomenclature to insist that Aristotelian matter has to be identified with localised particles. Although, of course, most moderns would comment that it is rather pointless worrying about what Aristotle would have thought about Maxwell's equations; we have moved on from Aristotle, and entered a realm beyond his wildest dreams.

7.3. Space and Time: from Aristotle to Newton

As a prelude to my discussion of Einstein's relativity, I need to address a subject which I have deferred, which is the vision of space and time developed by Galileo and Descartes, and adopted by Newton and all physicists until Einstein. I start with a brief summary of Aristotle's system which Galileo and Descartes overthrew. Space, for the Aristotelians, was a measure of extension and time was a measure of change. Without matter (and thus extension and change) there would be no space and time. The universe was seen as being ordered spherically, with matter either naturally tending to the centre of the sphere, or away from the centre, or (in the case of the astronomical bodies) given to circular motion around the centre (or an epicycle). The centre of the universe thus provides an origin against which every location can be measured; though we cannot say an origin of a coordinate system since Aristotle had no concept of coordinates or their uses.

Obviously, the Galileo/Cartesian system is very different. Because Galileo and Descartes wanted to combine geometry with physics, they obviously needed to relate the geometric concept of space and distance with something in the real world. In geometry, an Euclidean space is a region covered by various rules, such as parallel lines never meet and all non-parallel lines (in two dimensions) must eventually meet, one can construct straight lines and extend them indefinitely, circles can be generated by rotating a line segment around an end-point of that line, and that right angles are preserved under rotations and translations. On top of this, we can add the concept of distance as providing the overall scale of measurements, where a line segment of a certain distance can be mapped onto any other line segment of the same distance by a rotation and translation. Euclid's geometry required a background space on which the various line segments, circles and other shapes could be drawn, and

Galileo and Descartes' main observation (though, of course, building on the work of their predecessors going back to Bradwardine) was that the physical space could be represented by an Euclidean geometrical space. They differed in how this could be done. Descartes thought that physical space was identical to a Euclidean geometrical space; others said (more accurately in my view) that physical space could be mapped to a Euclidean space.

A mapping is a mathematical operation where two sets are compared. A set is a collection of objects. So, for example, we can treat $\{bananas,apples,plums\}$ as a set f and $\{yellow,green,purple\}$ as the members of a set c. The members of a set don't have to be numbers, though they can be. Sets can contain a finite number of members or an infinite number of members, they can be discrete or continuous, although obviously discrete finite sets are easiest to visualise. A mapping relates one member of a set to a member of another set. So, for example, we can introduce a mapping M which relates the set c with the set f, and denote it as $M : f \to c$. In this case, for example, the mapping might relate yellow with banana, green with apples, and purple with plums, so we can write $M : banana \to yellow$. We could also, of course, since we are just discussing things at an abstract level, introduce another mapping M' such that $M' : banana \to purple$. To say that M has a real-world application and M' doesn't requires us to identify these adjectives and fruits with real-world properties and objects, and then use observation to determine which of the mappings are worth pursuing. Sets can be constructed either from collections of physical objects or collections of abstract concepts, and a mapping can be used to relate the physical reality to the abstract description and vice versa.

A mathematical function is, in fact, just a mapping of this sort. We can define x as a set containing all real numbers, and y as another set also containing all real numbers, but distinct from x. The function $y = f(x)$ is then just a mapping: it associates each element of x with one particular element of y. We can write it as the set of all possible relationships $f : x_0 \to y_0; f : x_1 \to y_1$ and so on for every possible value of x_0 and x_1 (i.e. every possible number). Since writing this out for every possible number would require an uncountable infinite number of terms, this is obviously highly impractical, and the functional notation is easier, but the basic underlying principle is exactly the same. Mappings also have importance in other areas of mathematics, such as group theory and they are particularly crucial in topology, and these mathematical areas play very important roles in modern theoretical physics.

Mappings need not be reversible. For example, we have a mapping from *banana* to *yellow*, but we cannot apply a reverse mapping from *yellow* to *banana* because bananas aren't the only things that are yellow. There are numerous yellow objects which can all be mapped to that particular colour, and thus if we are given *yellow* we can't tell if it is a banana or a Korean inner city circular route bus. However, there is an alternative which is reversible, which is the one-to-one mapping. For example, I could relate the physical object of one particular cow with the description *an animal with the following DNA sequence* .... and as long as I get my DNA sequencing right and that cow doesn't have any twins or clones, then there is an exact mapping between the cow, a member of the set of animals that do or could walk this earth, and the DNA sequence, which is a member of the set of all possible DNA codes. The first of these is a physical object; the second a symbolic description, but the mapping allows me to relate one of them to the other and then back again.

Galileo's idea (although obviously he did not use this language since the mathematics was formalised after his time) was that we could construct one set $\mathbb{P}$ containing all physical points in space, and another set $\mathbb{E}^3$ containing all points in Euclidean space. He used Euclidean geometry because at that time it was the only one known. He then proposed that there was a one to one mapping between $\mathbb{P}$ and $\mathbb{E}^3$, that is to say that every physical point could be uniquely related to a geometrical point (in this particular realisation of Euclidean space), and every geometrical point could be uniquely related to a physical point. This doesn't tell us what physical space fundamentally is, but it does narrow down

the possibilities, and lists various properties that it must have for this identification to be possible. Physical space must be infinite in extent; it must have no determined centre or origin. It must be possible to identify points, and draw straight lines between them. It must be possible to perform the operations of rotation and translation. The space must be continuous: it is not possible to deform it so that two disjointed lines become a single line, and deformations of the space don't change the fact that lines intersect. Parallel lines extending infinitely must never meet; all right angles must remain right angles when translated or rotated, and so on. Obviously, there seems to be a great deal of correspondence between those of these claims that we can test through observation, and what we actually do observe. We cannot test that the universe is infinite, or that observed parallel lines never meet, simply because we cannot look that far. But the properties of Euclidean space that we can test did seem to be in good agreement with our everyday observations.

This, of course, is completely inconsistent with Aristotle's view of the universe. There is no centre of the universe to which things can naturally descend. There is no reason why circular motion should be preferred. Since the mapping between $\mathbb{P}$ and $\mathbb{E}^3$ makes no mention of matter, space would still exist even in the absence of matter. It cannot therefore be the measure of extension, or extension itself, but instead something which extended objects occupy. Descartes then provided the formalism which could be used to treat Euclidean space algebraically, and we were then all set up to treat space mathematically.

As was realized by Isaac Barrow, who occupied Newton's chair immediately before him and was one of Newton's teachers, time could be treated in a similar way, not as a three dimensional Euclidean space, but as a one dimensional Euclidean line. We thus have an understanding of space and time: a set of locations and instants which has a one to one mapping to a $\mathbb{E}^3 \otimes \mathbb{E}^1$ geometrical space. In particular the mapping of points in physical space to the $\mathbb{E}^3$ geometry is independent of time: the points don't move around as time progresses. This block universe view of space and time does not imply that time, the passing of time, or change are illusions. Rather, the reality of time is essential to it; the passing of time is represented by movement along the line, and change is measured by the differential operator. Clearly, this doesn't tell us what space and time fundamentally are, but it does (if true) rule out numerous possibilities and allows us to identify some of the features of space and time. In particular, it implies that time extends infinitely into the past and infinitely into the future. This is obviously something of a problem for the Christian who believes that the universe was created at a particular moment by God. However, it is not so much of a problem as might be perceived. For one thing, it was a familiar idea: Aristotle also proposed an eternal universe; Aquinas famously stated that the doctrine of a creation in time could not be proved by reason, but only had to be accepted through revelation (Aquinas, 1274, I.46)[10]. However, there were two ways around this: firstly, to say that God created all matter in the universe at a particular moment of time, or to remark that the idea that time extended infinitely into the past was not an experimental observation, but only a mathematical convenience. There is no fundamental reason why (for example) the universe could not have a first time, and be mapped to only (for example) a line segment in Euclidean space rather than the whole line, and thus time could have both a beginning and an end. Of course, this beginning and the end were only seen as artificial fixes put in by God; Newton's laws made no mention of them, and work for an infinite expanse of time. However the idea that time is truly something that could be mapped to an Euclidean line infiltrated the intellectual world, and by the early twentieth century few except orthodox theologians doubted that the universe was eternal.

[10]Note that Aquinas' cosmological argument was not that the universe must have a beginning (in time), but that there could not be an infinite sequence of instrumental (efficient or otherwise) causes (Aquinas' five ways contain one argument for each of the four causes, with the third way based on a second way of expressing efficient causality); this series of causes makes no mention of time, and the time between one cause and another is not mentioned nor needed, thus his argument is not inconsistent with an eternal universe.

This idea of time, of course, influences our ideas about past, present and future. We intuitively think that only the present in some sense really exists (and this, known as presentism, is the dominant view in Aristotelian philosophy). The past was existent; the future will exist, but only the present now exists. The idea of time as something that has a reversible mapping to an Euclidean line differs from this. The past, present and future are all equally real and existent. The only reason we perceive of the present as anything special is because our minds and senses are only aware of those things immediately local to us – both in space and time. For example, when we see a distant star, we only see the source of the light indirectly; what directly influences us is the photon of light impacting our retina. Thus our perception of time and space are both equally localised. However, we have no difficultly in believing that objects a long spatial distance from us are existent as we are, rather than objects coming into existence as they approach us and then going out of existence as they move away. The other difference between time and space is that time seems to flow in only one direction, while in space we can only move forwards and backwards. But this is again just a matter of the limitations of our perceptions. There is a direction in time; I would identify this with the direction of causal relations (i.e. the difference between efficient and final causes, which I take to be more fundamental than physics), while others would point to thermodynamics as the key (thermodynamics is the only part of the current expression of physical law which is not time reversible[11]). The point is that our present state is determined by the past. If we accept Newton's determinism (which I don't, but a similar argument can be constructed which doesn't depend on determinism; since I am discussing classical physics here, I will give the classical argument), then the past completely determines the present. Everything we are now is precisely determined by the state of the universe in the past, in particular our mental state at time t_0 (which might refer to the present, or to some other time; the point with this view is that there is nothing ultimately special about the present or the word *now*) is determined by the neural processes we had at times less than t_0, our initial formation at an earlier time, and the external sensations that have impacted us at earlier times. In other words, these will be imprinted on our brain, and we can thus look back at them, and for this reason at time t_0 we regard them as being in the past. But when we discuss the past, all we mean is *This thought is a space time event that can be mapped to the Euclidean line at time t_0, and these things happened at some time which can be mapped to points on the line earlier than t_0*. The future, on the other hand, is distinguished from the past simply by the arrow of causality. We were caused by events that happened in the past; the future is caused by us. But why shouldn't it be the future that caused us, and the past that we will cause? If so, then we will have no way of telling the difference from what we observe now; it would not have any effect on our observations, and only our labels would be modified. So time is not in a state of coming to be, being, and ceasing to be, but instead just a point on a line all of which is. We only see one particular point (the present) as special because each of our thoughts is an event which has to take place at a particular point along that line. That thought, that individual act of self-awareness thus artificially makes one point special and the origin of its own arbitrary choice of coordinate system. Things are categorised in relation to that thought.

I have made use of the word *event*, so I ought to describe what I mean by that. Euclidean space is constructed of a continuous set of points (this is a necessary prerequisite for Euclid's axioms), and we can, using Cartesian coordinates, map each point in the Euclidean space that physical space and time to four numbers, three to describe the spatial location, and one to describe the time. We describe this as a point. We can map this point to an instant and location in physical space/time, and the corresponding object in physical space/time is an event. We often think of events as things like collisions or trains leaving stations, but in this notation the event just describes when and where such things can happen.

[11]More technically, physical law is symmetric under the simultaneous application of time reversal, parity and charge conjugation.

There are two further ideas which are important to Newton's idea of space and time. What I have been discussing so far is an $\mathbb{E}^3 \otimes \mathbb{E}^1$ manifold which provides a background to all of physics. This is known as absolute space and time. It contains no origin or preferred coordinate system; although we apply a coordinate system to it to enable our calculations, this choice is wholly arbitrary and the physics does not depend on it in any way. A coordinate system provides a further mapping, this time from Euclidean space-time to the set of real numbers, $\mathbb{E}^3 \otimes \mathbb{E}^1 \to \mathbb{R}^3 \otimes \mathbb{R}^1$. We should not think that because of this mapping that Euclidean space is equivalent to the coordinate system we apply to it (nor should we think that real space and euclidean space are equivalent). They are not. In particular, as soon as we choose a coordinate system we introduce a special point as the origin and special reference frame which is not present in Euclidean space and even less so in the real physical space. Physics relates to real space via Euclidean space, and not to the space of real numbers. Thus physics ought to be independent of any arbitrary choices we make when performing these mappings, including the choice of coordinate system. We should also not make the mistake of thinking that the $\mathbb{E}^3 \otimes \mathbb{E}^1$ geometrical space encapsulates everything about real space. It certainly describes some aspects of real space, because otherwise modern physics would not work at all, but not necessarily everything and it might introduce some freedom which does not apply to reality.

In addition to this absolute space, we have relative space, and also the corresponding relative motion. There is also a similar concept of relative time. Relative space is space after we have chosen some particular coordinate system. It measures distances and angles with respect to some particular point, and most usually distances between two bodies. As far as linear motion is concerned, we can only observe relative motion and relative distance and relative duration. Why then postulate the absolute motion, time, and space? Firstly because relative location is not unique: there are an infinite number of ways in which it can be defined, depending on which pairs of particles we use as the 'standard' distance. Absolute space, however, is unique, and can be used to construct relative space. Newton's mechanics does not make sense unless there is some absolute space underlying everything.

We can say that two bodies are at rest with respect to each other if the distance separating them does not change in time. Similarly, we can measure relative linear velocity. If the distance between the two bodies is increasing or decreasing with respect to time, then the relative velocity is the rate of this increase or decrease. If we introduce a third body, then we have another two sets of relative velocities. Relative locations and velocities are related to each other, in Newton and Galileo's system, by simple addition. So if $\vec{\mathbf{d}}_{AB}$ is the vector describing the separation between the bodies A and B with $\vec{\mathbf{v}}_{AB} = \frac{d\vec{\mathbf{d}}_{AB}}{dt}$ the relative velocity, and similarly $\vec{\mathbf{d}}_{AC}$ and $\vec{\mathbf{v}}_{AC}$ the relative distances and velocity between A and C, then we can compute the distance and velocity between C and B as $\vec{\mathbf{d}}_{BC} = \vec{\mathbf{d}}_{AC} - \vec{\mathbf{d}}_{AB}$ and $\vec{\mathbf{v}}_{BC} = \vec{\mathbf{v}}_{AC} - \vec{\mathbf{v}}_{AB}$. These are just the expressions expected in a Euclidean space with all observers measuring time in the same way. One of the consequences of Newtonian and Galilean relativity is that as far as straight line motion is concerned, one cannot tell whether or not a body is at rest with respect to absolute space or not. This is a major difference in comparison to Aristotle's physics.

There is one type of motion, however, where we can observe whether we are at rest with respect to absolute space, and that is rotational motion. The centrifugal forces experienced by a rotating body can be measured, and these are zero when the body is not rotating absolutely. This is because rotational motion implies a continual change in velocity, or acceleration, and we can measure absolute acceleration: they are just the same as relative acceleration.

To see this, we need to introduce the concept of inertial frames. An inertial frame is described by a coordinate system whose origin moves with a constant velocity with respect to absolute space. For example, suppose somebody sitting on a train moving at a constant speed down a straight track creates a coordinate system with himself as the origin. He will

measure the passenger in the next seat as being at rest with respect to him. His frame of reference gives one inertial frame. Outside the train, we have a man driving a Jeep across a the field. Once again, he can construct a coordinate system with himself as the origin. Again, he travels at a constant velocity. Finally, we can have a man sitting on a park bench, who once again develops his own coordinate system and inertial frame. All of these see the other one as travelling at a constant velocity with respect to each other, but (for example) the man on the bench measures the man on the train as travelling at one velocity, while the man on the Jeep measures him with a different velocity, while the man on the train believes himself to be at rest with respect to his coordinate system. There is, of course, no contradiction here, because they each use different ways of measuring speed. There would only be a contradiction if two people in the same inertial frame made a different measurement of velocity.

In symbols, we can write the coordinates used by the man on the bench as $(\vec{\mathbf{x}}_B, t_B)$. He measures the Jeep as travelling at a velocity $\vec{\mathbf{v}}_{BJ}$ and the train as travelling with a velocity $\vec{\mathbf{v}}_{BT}$. It is necessary to specify both the bench and the Jeep or train when constructing a notation for the velocity because we need to know in which inertial frame the velocity was measured. Now suppose that the coordinates used by the man on the Jeep are $(\vec{\mathbf{x}}_J, t_J)$ and those used by the train passenger $(\vec{\mathbf{x}}_T, t_T)$. We can then write down equations to convert from one coordinate system to another:

$$\begin{aligned} t_B &= t_J = t_T \\ \vec{\mathbf{x}}_T &= \vec{\mathbf{x}_B} + \vec{\mathbf{v}}_{BT} t_B + \vec{\mathbf{x}}_{T0} \\ \vec{\mathbf{x}}_J &= \vec{\mathbf{x}_B} + \vec{\mathbf{v}}_{BJ} t_B + \vec{\mathbf{x}}_{J0} \end{aligned} \tag{228.1}$$

$\vec{\mathbf{x}}_{T0}$ and $\vec{\mathbf{x}}_{J0}$ are vectors which we need because the origins of the various coordinate systems might be different. These are constant in time. One also does not need each observer to have the same origin for time, but for simplicity I have assumed that the bench man's $t = 0$ is the same as the Jeep man's $t = 0$ and the train man's; this is easily corrected by adding the appropriate constants to the first of these conditions, but makes no difference to my overall conclusions so for simplicity I have neglected this. The second thing that I have assumed is that each observer has agreed to use the same length scale, for they are all measuring distances in metres or all in inches, and not some in meters and some in inches. If they are using different length scales, then all we have to do is divide each $\vec{\mathbf{x}}$ and $\vec{\mathbf{v}}$ by some fixed distance that all observers can agree to adopt as a standard, which they each measure in their own units. For example, they can each measure the height of a particular post or the length of the train, and divide all the distances they measure by what they measured for this object. A similar consideration applies to the units used to measure duration: if they differ from one observer to another, we need to divide each variable representing time by some standard duration agreed on by each observer – maybe the Jeep gives one blast of its horn, and they all time how long that takes. Individual measurements of distance and time are convention dependent[12]. Ratios of distances or durations, as long as consistent conventions are used, are the same no matter what conventions are chosen. In this way, each distance measurement or time measurement is expressed in a way that is independent of the arbitrary choice of how long a meter is, and because there is no longer any arbitrariness in the equations, all of our observers should agree. Presumably these standard distances and durations (i.e. the height of the post, or duration that the Jeep's horn sounded) will not depend on the relative velocities between the observers or on time; if so then this need to standardise units makes no significant difference to the rest of the argument – it just adds a few more symbols to the equations – so I may as well keep things simple and leave this complication out by assuming that all observers are using the same scales for distance and time measurements.

[12]A convention being something which we can arbitrarily choose, and must arbitrarily choose to make any measurement, such as how long a meter is.

What do these equations tell us? Suppose that there is an object, which all three of them see, maybe a tree or a plane flying overhead – this time it doesn't have to be at constant velocity. Each observer maps out a path for that object in their own coordinate system, so, for example, the observer on the bench uses his own $\mathbf{x}_B(t_B)$ as a variable which describes the path of this object. The observer on the train also measures the path of the object using his coordinate system and comes up with $\mathbf{x}_T(t_T)$. Clearly these two functions are not going to be the same, because they were measured in different coordinate systems. But equally clearly they are going to be related in some way, because they describe the same physical object and ultimately the same physics. The equations above describe that relationship; how to convert from one coordinate system to another.

You might ask what is so special about inertial frames? Why not consider coordinate systems where the observers are accelerating relative to each other, or accelerating their acceleration, or with no easily categorised relative motion of any sort? Inertial frames have the advantage that the functions mapping from one coordinate system to another are linear functions of the distances and durations – they only depend on constants and single powers of the distance and time, and not time squared or distance squared or distance times time (indeed, I will use this property to define what I mean by a set of inertial frames). Thus the mathematics is much simpler. The second issue is that Galilean relativity (and thus Newton's mechanics) stops working when we start considering anything beyond inertial frames (which is slightly discouraging). So, since for this discussion we are playing the role of devoted disciples of Newton, we will sweep this question under the carpet for now, put our fingers in our ears, and forget that anyone asked it. But it is an important question, and we can't ignore it forever. It will resurface when we discuss the general theory of relativity.

This Galilean relativity is very different from philosophical relativism, which states that 'truth' depends on the observer and is not objective. The main difference is that coordinate systems are a tool used by physicists but not a part of physical reality. So even though each observer comes up with a different function to describe the path of the object they are tracking, these functions are merely descriptions, not reality. In each case the reality is the same. We map reality to a mathematical abstraction, in this case a coordinate system, perform calculations, and then map the abstraction back to reality. Our abstractions depend on various arbitrary choices we make or conventions we take and are (in a limited sense) observer dependent; reality does not depend on these conventions, and neither does our final result when we map our abstraction back to reality; this last step removes any convention dependence we introduced earlier. The philosophical relativist does not demand consistency; he accepts all opinions regardless of whether they can be mapped onto each other. For example, in moral relativism, one person might say that infanticide is sometimes morally good, and another that it is always a moral evil, and both opinions are deemed valid even though they are in direct contradiction with each other. In physical relativity the two measurements are identical in meaning but only seem different because they use different languages. The analogue is one ethicist who says 'Infanticide is always morally wrong' and the other who says 'Infanticide is never morally acceptable'; different words conveying the same idea. If the relativist tries to use Galilean relativity (or Einstein's relativity, which is based on the same underlying principle, just with a different formula because it doesn't use Euclidean geometry) to justify their philosophy, then they are making the mistake of confusing the abstraction with the real world.

But it is even worse than this. A competent physicist sitting on the bench would never state *The path of the aeroplane is described by the function* $\boldsymbol{x}_B(t_B)$. That statement is incomplete and consequently of no use to anyone. You state your conventions along with the result. The physicist would instead say *The path of the aeroplane is described by the function* $\boldsymbol{x}_B(t_B)$ *in this specified coordinate system where I am always at the origin.* That statement is complete, and objective. The observer in the train and the Jeep would make their own measurements, calculate the path of the plane in the coordinate system most natural to

them, but they will know how to convert their result from their coordinate system to the bench coordinate system for comparison, and when they do so they will get precisely the same result, to within the various experimental imprecision.

Equation (228.1) doesn't just tell us how to convert distances from one coordinate system to another, but also velocities and accelerations. We just have to differentiate with respect to time. The first differential gives us the rule for adding velocities measured by observers travelling at different speeds (given that $t_T = t_B = t$):

$$\frac{d\vec{\mathbf{x}}_T}{dt} = \frac{d\vec{\mathbf{x}}_B}{dt} + \vec{\mathbf{v}}_{BT}. \tag{230.1}$$

This just states that the velocity of the aeroplane measured by the man on the train is just the velocity of the aeroplane measured by the man on the bench plus the relative velocity between the bench and the train. Which is a bit of a mouthful to write in words, but all very natural and in line with common experience. The second differential gives us the rule to match the accelerations:

$$\frac{d^2\vec{\mathbf{x}}_T}{dt^2} = \frac{d^2\vec{\mathbf{x}}_B}{dt^2}. \tag{230.2}$$

What this tells us is that in every inertial frame, the same acceleration is measured. This means (assuming that mass is independent of the inertial frame) that the same forces are measured in every inertial frame. Since Newton's laws basically say that physics is ultimately reducible to forces acting on matter (the field picture has the fields creating forces which act on matter, so the same argument should apply after Maxwell and Faraday re-expressed mechanics in terms of fields), this means that the same physics applies in every inertial frame. This is good. The choice of inertial frame is just a choice of coordinate system, and the choice of coordinate system is entirely arbitrary. It is a matter of our own conventions rather than the real world, and thus real world physics shouldn't depend on it. That means that each observer measures the same gravitational, impact, electrical and magnetic forces no matter how fast they are travelling or in what direction. Thus the disciples of Galileo and Newton, perfectly naturally, concluded.

But there is a problem: Maxwell's equations. Maxwell's equations state that the speed of light in a vacuum is calculable from the the parameters ϵ_0 and μ_0, which describe the strength of the electric and magnetic forces. From what we have just seen, the strength of the forces should be independent of the inertial frame. However, we have also seen that in Galilean relativity velocity (including the speed of light, $c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$) is different in different inertial frames. There is a contradiction here, and something must be wrong. It looks as though we have to choose between the principle that physics is independent of the inertial frame on the one hand, and Maxwell's equations and in particular the connection between electromagnetism and optics on the other. But the principle is one of the foundational axioms of physics – it is derived from the contention that there are universal objective laws of physics, and if that's not true we may as well give up and go home because we are never going to make progress. Maxwell's equations are supported by a huge amount of experimental evidence. So we have an apparent contradiction, which means that one of our assumptions must be wrong. So are there any other options? Any other assumptions besides Maxwell's laws and the coordinate invariance of the laws of physics? Well, we did assume that mass and electric charge were independent of inertial frame, and that might be worth looking into. Equally we assumed that the unit used by each observer to measure distance (and also duration) did not dependent on the relative velocities between the observers, and that too needs re-examining. But the most important mistake in our assumptions is far more fundamental than any of these. We assumed that the universe can be mapped to an $\mathbb{E}^3 \otimes \mathbb{E}^1$ Euclidean geometry. It turns out that this is ultimately what caused the contradiction. There is a mapping between physical space/time location and a geometrical abstraction, but that geometry is not Euclidean.

7.4. The Formalisation of Mathematics

The nineteenth century saw great advances in diverse areas of mathematics; but one thing that all these advances had in common was that they made mathematics more abstract, and more divorced from observation. Mathematics started off as the investigation of number, which was needed because we observe number in the physical world and needed to know how to combine and divide them. Geometry came next, which is the study of shape, but again, this is strongly related to observation, because we see shapes in nature, and it is just a short step from this to the idealisations used by the geometers. With algebra and later calculus, we are becoming more abstract; but still not wholly abstract. Differentiation, for example, was used to understand velocity and acceleration, which are again observed quantities. Integration is used to understand area and volume. Thus we are still very much tied to the physical observed world; the symbols may be abstractions, but they are tied to things we can experience. In all these areas, historically, the mathematics was developed in order to properly express advances in what we now call science.

Set theory, however, which some people take to be the underlying foundation beneath all other areas of mathematics, is different. Here the formalisation of mathematics is complete. The set theoretician does not care what the sets represent - it makes no difference to him. Instead, he merely studies the possible relationships between the different sets. In other words, mathematics became entirely abstract and formal; shedding its dependence on the physical world and in particular our common sense derived from the physical world. That doesn't mean that it is no longer applicable to the physical world - far from it: modern twentieth century physics found itself reliant on the abstract mathematical techniques developed from the nineteenth century; almost everything the mathematicians worked on was later found to play an essential role in some part of physics.

Instead of using physical justification for their work, the mathematicians turned to self-consistency as their principal guide. They would select various axioms, true statements that cannot be proven as a consequence of the calculations, and then build up the various consequences that follow from those axioms. From a small number of basic principles, a huge number of conclusions follow. The motivation for this approach was Euclid's geometry, but Euclid's axioms are not the only ones which could be adopted. Indeed, mathematicians delight in trying out different sets of axioms and seeing where they lead.

We cannot eliminate the need for axioms – that has been formally proven; but we can take them from intuitive principles. For example the principle of non-contradiction is a basic axiom of metaphysics, logic and mathematics. It cannot be proven, but that doesn't mean that it isn't true. However, axioms can be tested by comparing the conclusions against observations of nature.

This shift of mathematics from the concrete to the wholly abstract proved to be incredibly prolific. It allowed new branches of mathematics to emerge - branches which would prove vital in understanding physics. But the very fact that we can do it points to something beyond physicality. There is no question that the mathematicians are studying truth, but truth divorced from the physical world. The shift in focus turns us away from empiricist presuppositions into supposing a system where there is an abstract and non material structure underlying the universe (since the mathematics is applicable to the study of physics).

Equally, the link between mathematics and logic was strengthened in this time period; the nineteenth century saw the revival of logic as a major topic and tool of study, only this time it was symbolic and abstract, taking the language and tools of set theory. Propositions are written in terms of underlying symbols; it does not really matter what those symbols represent, since the same manipulations can be applied regardless of those meanings. For example, if a represents Socrates, P is the set of all men and M the set of all mortal beings, then the relation Socrates is a man is reduced to $a \in P$, while the proposition that all men are mortal reduces to $P \subset M$ (the set P is a subset of the set M), and the rules of logic then tell us that $a \in M$, i.e. that Socrates is a member of the set of all mortal beings. But

equally, we could use a to represent Fido, P to represent the set of all possible dogs, and M to represent the set of all possible mammals. Indeed, we can proceed with the logical argumentation without first specifying what a, P and M represent. In this way, logic again becomes wholly abstract in its execution, though still applicable to real world situations.

Mathematical truth is a consequence of the axiomatic method, which proceeds by establishing first correct propositions formed from symbols representing objects and sets and relations between them. A propositional calculus then allows the combination of these propositions according to the logical rules. We can then assign a truth value to each proposition, which can be denoted by the symbols 1 for true and 0 for false. Finally, we take a certain subset of propositions, and declare them to be true: these are the axioms of the system. If the axioms are chosen well, then from a small number of axioms, we can determine the truth or falsity of every other proposition in the system. If a proposition is proved to be true as a consequence of the axioms, it is known as a theorem, and added to the account of true statements, and can be used in subsequent proofs.

The choice of axioms is not arbitrary: one seeks to prove the most from the fewest axioms. Introducing a new axiom randomly is likely to lead nowhere; to be adopted a mathematical idea needs to be useful in that it does not contradict established truth and leads to new ideas. But the consequence of this method is that mathematics is continually expanding, as new truths are proven. We do not know what the limits of the mathematical method are; we do not know if it has limits.

Mathematics has been reduced to a science of relations. Its underlying structure is entirely abstract; one does not need to specify what is being related to what. It does not have its own meaning. But instead, it can be given a meaning when it is applied to a particular field of study. For example, in physics, we assign various mathematical structures to represent various physical objects, and then the mathematics takes over by showing the relationships between those objects so represented. Of course, this means that what we deal with in physics are only representations of the real objects, not the physical objects themselves. We capture part of their nature in these representations; there is no reason to suppose that we capture the whole of their nature. But that the abstract language of mathematics is applicable to physics, with that application being successful, implies that, in part, physical objects can be represented abstractly. We have to ask why this process of abstraction works; and the most obvious suggestion is that the representations used in physics correspond to something that is truly present in the physical object. In short, that there is an underlying immaterial layer to every material object. The representation cannot capture everything about the physical object: there must be something not expressible mathematically from which to formulate the underlying axioms; since mathematics without context is meaningless, if physical objects were entirely mathematical then they too would be devoid of their concrete existence. There must therefore be part of a physical object which cannot be represented in mathematical terms: which can be thought of as the underlying matter; equally there is a structure which shapes this matter which can be represented mathematically, and this structure is what is studied in physics. In addition to the matter, there is a concept underlying the substance which we can understand intellectually.

We conclude that there is a formal and immaterial structure underlying the material world and part of material substances. This structure is what is uncovered by modern physics. In Aristotelian terms, the mathematical physics is not the study of matter, but of form.

7.5. Special Relativity

Euclid wrote his major work on geometry around 300BC, and it is a testament to his genius that no alternatives were found for over two thousand years; there were many great minds studying geometry in that period, and yet (aside from Descartes' introduction of coordinate spaces) their work just amount to footnotes to the *Elements*. It took another

genius to show that Euclidean geometry was not the only alternative, and that man was Carl Friedrich Gauss, one of the greatest mathematicians there has ever been. We have have already encountered him as the originator of one of Maxwell's equations, but he had many contributions in various areas of mathematics. Subsequent to Gauss's work, others such as Bolyai, Lobachevsky, and most importantly Bernhard Riemann extended his work during the course of the nineteenth century.

Euclid wrote his geometry based on five axioms; from these he proved everything else. The first four axioms are

(1) A straight line segment can be drawn joining any two points.
(2) Any straight line segment can be extended indefinitely in a straight line.
(3) Given any straight line segment, a circle can be drawn having the segment as radius and one endpoint as centre.
(4) All right angles can be mapped onto each other through a sequence of transformations that preserve length (such as rotations and translations in Euclidean space).

There was, however, for a long time a debate about whether the fifth axiom, that all lines that are not parallel eventually meet, could be derived from the previous four. Gauss showed that his new hyperbolic geometry satisfied Euclid's first four axioms, but not the fifth, thus demonstrating that the fifth Euclidean axiom could not be derived from the first four. Special relativity is basically the idea that the geometry of the universe is not Euclidean but hyperbolic; although Einstein did not realise this at first until it was pointed out by Einstein's old supervisor, Herman Minkowski (for an introduction to special relativity, see (Einstein, 1920; Wald, 1992; Taylor and Wheeler, 1992)).

Firstly, we need to go into a little more detail into geometry. A geometrical system is uniquely defined by a number of properties. The first is its topology, which is related to the continuity of the structure, telling us which translations are continuous and which are discontinuous, i.e. where a small jump in the location leads to a large jump in the coordinates. For example, we could define a two dimensional system with just one coordinate. We can take our two Cartesian coordinates, and splice them together into one number (the even digits give the x coordinate while the odd digits give the y coordinate, so, for example, $x = 12.345$ and $y = 678.9067$ leads to the number 61728.39405607). However, this isn't much use because a continuous change in the space (x increasing by a small amount while y remaining the same) can lead to a discontinuous change in the coordinate. The topology also informs us of any periodicity in the system (for example, telling us that an angle of 360° is equivalent to an angle of 0°), and the dimensions of the system. Next we need an affine structure, which tells us how to define a straight line. Finally, we need a metrical structure, which is how we define distances. The metric is of particular interest, because it plays a key role in both special and general relativity. The metric can be represented by a matrix, which I will denote as g_{ij}, which tells us how individual differences in each direction add up to a squared length or distance. So, for example, we can have a vector $\vec{\mathbf{d}}$ which represents the distance between two points. Pythagoras tells us how to construct the distance element s associated with this vector in Euclidean space: $s^2 = d_1^2 + d_2^2 + d_3^2$, where d_i are the individual components of $\vec{\mathbf{d}}$. This can also be written in terms of a matrix/vector multiplication, where we write

$$s^2 = \sum_{ij} d_i^T g_{ij} d_j, \tag{233.1}$$

or just, using a convention of Einstein which was basically 'I'm too lazy to keep writing summation signs all over the place',

$$s^2 = d_i^T g_{ij} d_j. \tag{233.2}$$

The symbol T means that we are taking the transpose of the vector, i.e. converting three numbers in a row into three numbers in a column. In this notation it doesn't really do very

much, but it is worth putting it in there to remind us of what to do if we want to write the vectors out in full. Something with one index is a vector, and something with two indices is a matrix. In this case we are talking about three dimensional space, so g_{ij} is a 3×3 matrix. So this prescription states *take your vector* $\vec{\mathbf{d}}$*, multiply it by the matrix* g*, and then take the inner product with* d *again and you get the square of the distance of* $\vec{\mathbf{d}}$. In Euclidean space time, we know that the distance is just the squares of the elements of the vector, which means that g is just the identity matrix, with $g_{ij} = 1$ when $i = j$ and zero elsewhere. So what's the point of putting it in there when it doesn't actually do anything? We do actually see something vaguely interesting. We know that a rotation preserves distance. The rotation of a coordinate vector can be described using a matrix R_{ij}, where the rotated vector $\vec{\mathbf{d}}'$ satisfies $d'_i = R_{ij}d_j$. This means that the distance relation in the rotated system is $s^2 = d'^T_i g_{ij} d'_j$, but this is also $s^2 = d^T_k R^T_{ki} g_{ij} R_{jm} d_m$ and simultaneously, since distance is preserved, $s^2 = d^T_i g_{ij} d_j$. This implies that $R^T g R = g$. This is actually a well known property of the rotation matrices, where g is just the identity operator.

Hyperbolic geometry makes one small change from Euclidean geometry. Instead of having g as the identity operator, we change one of the diagonal elements into -1, thus

$$g = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{234.1}$$

This might only seem like a small change, but it is quite profound. Euclid's first four axioms continue to be satisfied; this change in how we measure distance is the only change we are making to the geometry. The new geometry is fully self consistent, only now we have sent Pythagoras packing. The distance element now reads $s^2 = -d_1^2 + d_2^2 + d_3^2$. Non-parallel flat planes (since we are in three dimensions) need no longer meet. Of course, this means that rotations no longer conserve distance, except those whose axis is parallel to the direction represented by the coordinate d_1. We can write a rotation matrix in terms of the (hopefully) familiar trigonometric functions sin and cos

$$R = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{234.2}$$

where θ is the rotation angle, and this satisfies $R^T R = 1$ since $\cos^2\theta + \sin^2\theta = 1$ (which is basically just a restatement of Pythagoras' theorem). But $R^T g R \neq g$ for our new metric. Instead, we need to apply the hyperbolic rotation

$$\Lambda = \begin{pmatrix} \cosh\theta & \sinh\theta & 0 \\ -\sinh\theta & \cosh\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{234.3}$$

where cosh and sinh are functions which satisfy $\cosh^2\theta - \sinh^2\theta = 1$. This, plus another hyperbolic rotation in the $(1,3)$ plane, and a normal rotation in the $(2,3)$ plane are the three operations which leave the distance invariant.

Right, that's enough mathematics for the moment, so back to where we left off with the physics. We have a problem with Maxwell's equations, and that they seem to express a speed, which should depend on the inertial frame we are in, in terms of the strength of the electric and magnetic forces, which shouldn't. So something has gone wrong somewhere.

In the late nineteenth century, the nature of light was something of a puzzle. It was a wave, but in what medium? It was proposed that there was a fluid, not made of ordinary matter so we could pass through it without noticing it, which existed in all space and time. This was known as the aether. Light would then be waves in the aether, and very similar to waves on the surface of a body of water. In this case, there would be one preferred frame of reference: that in which the aether was on average at rest. In this case, one should be able to measure a tiny difference in the speed of light depending on whether one is travelling

along the direction of the aether or in another direction: its a bit like an ocean canoeist travelling with the waves or across them. If he travels with the waves, they will seem to come to him slower than if he is travelling perpendicular to them, since every time a wave reaches where he was, he has moved a little bit further. The earth is travelling with respect to the aether, so if you measure the speed of light in the direction of the earth's motion, you should (if this theory is correct) get a slightly different result compared to the direction perpendicular to the earth's motion. This difference in speed is, of course, tiny. It can be converted to a tiny difference in distance by multiplying by a short time element, but this seemingly makes the measurement problem equally hard. However the wave nature of light helps us here, because we can use interference effects to measure small distances. Take a beam of light, split it in two ways, reflect each beam off a mirror, recombine them, and look at the result. You will get either constructive or destructive interference depending on the difference in the two paths taken by the light. Adjust the equipment so that you get constructive interference. Now we come to the clever part. You physically rotate the equipment around the axis of one of the split beams. If light is travelling at different speeds in different directions, then the interference pattern should change. You have just proved Maxwell's equations wrong, and won yourself a Nobel prize.

But, of course, the experiment doesn't prove Maxwell wrong. The speed of light is the same in every direction, and in every inertial frame (it doesn't matter what time of day or year you conduct this experiment). Maxwell was right, which means that Galilean relativity is wrong.

Einstein thought about this problem, and realised that one issue was in how we measure a standard distance. You will recall that one aspect of Galilean relativity which I skipped through quickly was that we could only compare ratios of distances. In other words, each observer had to measure the length of some object in their own units so that they could make a comparison. But how do we make such a measurement? We can measure distance from us by reflecting light off it, and measuring the time taken for the light to return. However, if we are moving towards the object, it will seem as though it is slightly closer to us, because by the time the light has bounced back to us, we would have moved onwards a bit. This means that the length scale is slightly contracted in a moving frame.

Special relativity is primarily concerned with the relationship between coordinate systems in two different inertial frames. We label one set of coordinates in one inertial frame as (ct, x, y, z), and the other, in a different inertial frame, as (ct', x', y', z'), where c represents the speed of light. We know from the basic principles that led to Galilean relativity that there must be some way of converting from expressing measurements made in one coordinate system into the other coordinate system. We also know, from that Galilean relativity gets the transformation for objects travelling at the speed of light wrong, that Galileo's own equations need tweaking. The goal in special relativity is to find those tweaks, and then to re-construct Newton's dynamics and second law as necessary to maintain self-consistency. This means that we will need to have to consider how location, time, energy, momentum, force and the electromagnetic fields and potentials, and various other representations of physical quantities transform when we change the coordinate system.

If a light beam is emitted at the origin in space and time, then the path of light measured by an observer in one particular inertial frame is given by

$$c^2t^2 = x^2 + y^2 + z^2. \tag{235.1}$$

This is basically just the statement that the beams of light are measured as travelling at the speed of light conventionally denoted as c. The right hand side of this equation is the square of the distance travelled, the left hand side the square of the velocity times the time travelled. In an alternative inertial frame, this is,

$$c^2t'^2 = x'^2 + y'^2 + z'^2. \tag{235.2}$$

We will rotate the coordinate axes so that the second inertial frame is moving with a velocity v along the x axis with respect to the first inertial frame. We expect that we should be able to write the primed coordinates as a function of the unprimed coordinates. We have two conditions: firstly, that the above equations are satisfied, and secondly, that the results reduce to Galilean relativity in the limit that the speed of light is infinite. Why do we want this second condition? We know that Galilean relativity is satisfied incredibly well from our day-to-day observations. So whatever final result we get, it should be similar to what Galileo proposed. However, we also know that Galilean relativity breaks down at velocities close to the speed of light. Now, if the speed of light were infinite, then we could never get close to it, so Galilean relativity would work at any velocity. Of course, we can't change the speed of light in the physical world, but we can change it in our equations. It is, after all, just a number. If we want to know what our world is like, we must set that number to the measured value. But the theory allows us to ask an additional question: what would the universe be like if the speed of light were some other value (and everything else the same as much as is possible)? We can hypothesise that the world would be Galilean. So these are the equations that describe how to map from one coordinate system to another in special relativity, which provide the unique solution to these two conditions plus the condition that the new (primed) coordinates are linear functions of the old coordinates (which is our definition of an inertial frame):

$$\begin{aligned} z' &= z \\ y' &= y \\ x' &= \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(x - \frac{v}{c}ct) \\ ct' &= \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(ct - \frac{v}{c}x). \end{aligned} \tag{236.1}$$

I have multiplied the time variable by the speed of light everywhere so that it has units of distance. It can be seen that if $\frac{v}{c}$ is set to zero (while keeping v non-zero) we recover Galilean relativity (equation (228.1)), as desired. We can even go further than this. There is a mathematical procedure known as a Taylor expansion. Polynomial functions (the sum of integer powers of x multiplied by coefficients) are usually easier to deal with than more complicated functions such as square roots. However, for almost every function $f(x)$ – the exception is if the function contains a discontinuity – it is possible that 'near enough' to a point x_0, we can write $f(x)$ as $f(x_0)$ plus a polynomial of $x - x_0$. The polynomial has an infinite number of terms, but we can still deal with that in certain circumstances. What is meant by 'near enough' is a little less clear and depends on the function. But if $x - x_0$ is small enough, then the term in the expansion proportional to $(x - x_0)^2$ will be smaller than the term proportional to $(x - x_0)$, and the term proportional to $(x - x_0)^3$ smaller still, and so on continually up to infinite order. If we want to evaluate the function exactly, this isn't much good: we have to deal with an infinite number of terms, which is impossible. But if we only want to evaluate the function approximately, then we can use this, just discarding all the terms above a certain order. If we want to make it a bit more accurate, then we just add a few more terms. Why not just evaluate the function? I will just say here that, in particular if we need to multiply this function by some other function and ingrate the result or solve a differential equation, using a polynomial expansion is the only way we can get close to the actual result: we can't use the exact function because a closed form solution to the full equation is not known. However, in these cases, the solution provided by the controlled approximation of a truncated series expansion can be solved, and gives some answer. Since to have something is better than to have nothing, we have to make do with the approximate solution.

In physics, we often only need to calculate theoretical results approximately. Our goal is to compare theory with experiment, and every experimental measurement contains an error: it only gives a result up to a certain precision. Indeed, the initial conditions for any theoretical calculation are derived from experiment, which means that we are never going to get the exact result anyway: every theoretical result is also imprecise. So we don't need a perfect theoretical computation, we only need the theoretical errors to be much smaller than the experimental errors so they don't influence the final comparison. An exact calculation would be better, but this isn't always possible, or if possible, it might simply be too much effort in comparison to the approximate calculation. Series expansions are thus one of the primary tools used by theoretical physicists and phenomenological physicists (phenomenology is the branch of physics that compares theory with experiment). A series expansion also has the advantage that we can estimate the error to a reasonable accuracy: we know that each term is smaller than the previous one, so if we also know how much they decrease by, we can estimate with a good accuracy how large the correction from the next few terms will be.

In the case of the Lorentz transformations, the series expansion in small v/c tells us something else. We can write

$$\begin{aligned} x' =& (x - vt)(1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \ldots) \\ ct' =& (ct - \frac{v}{c}x)(1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \ldots). \end{aligned} \tag{237.1}$$

This tells us the relationship between special relativity and Newtonian physics. At the cost of a little over-simplification (since there is more to Newton's physics than just Galilean relativity), Newtonian physics is just the leading order term in the series expansion of special relativity physics in $\frac{v}{c}$ or the relative velocity of the inertial frame divided by the speed of light. This is why Newtonian physics works so well. It is the simplest, and most inaccurate, approximation we can make to the real physics. But because the expansion parameter $\frac{v}{c}$ is so small in our every-day experience, the correction is so much smaller than the usual experimental errors we encounter that we may as well just ignore it. Of course, if you have incredibly accurate experiments, or move very close to the speed of light, then you can't ignore it, and, of course, these experiments have been done to confirm that special relativity is correct (and it is correct as far as we can tell: Lorentz invariance, which is a symmetry derived from the Lorentz transformations, is a key element of both the standard model of particle physics and of the standard model of cosmology and has thus passed experiments to an incredible precision for everything from the smallest particles we know of to the scale of galaxies and the universe as a whole). But for day to day life, you can't tell the difference between Newton's physics and special relativity.

Notice something different between special relativity and Galilean relativity. In Galilean relativity, the time coordinate did not change from one inertial frame to another. In special relativity it does, and it also depends on the location in the original coordinates. There is thus an unexpected (to the Newtonian) mixing between space and time. For Galileo and Newton, and indeed everybody up to the early twentieth century, space and time were two entirely different things. Now we discover that they are not. One's perception of time depends on how one moves through space. What this suggests is that instead of our usual three dimensional vectors to denote the spatial location and having a separate time coordinate, we should group them together into a vector with four components. In this case the Lorentz transformation just becomes a matrix operation on this four dimensional vector.

Now we are in a position to get to the main point, and where the non-Euclidean geometry comes into view. We noted that $c^2t^2 - x^2 - y^2 - z^2$ is the same in every inertial coordinate system. This looks similar to Pythagoras' formula for the squared distance. In fact, we can write it in a vector matrix form with a metric tensor which is conventionally given

its own symbol because it is so important, $\eta_{\mu\nu}$, which is the Minkowski matrix, named after Einstein's former teacher who was the first to notice that special relativity had this geometrical interpretation. We will write the space time vector as $x_\mu = (ct, \vec{\mathbf{x}})$, where we use a Greek index to show that it is a four component vector, and we call time the zeroth dimension, so we can keep the spatial directions as dimensions one, two and three. $\eta_{\mu\nu}$ has $\eta_{00} = 1$, $\eta_{11} = -1$, $\eta_{22} = -1$, $\eta_{33} = -1$, and all other components zero. η is nothing but the metric tensor of a $3+1$ dimensional hyperbolic geometry. What we have been doing is writing location and time in terms of a hyperbolic rather than Euclidean coordinate system.

We can then write that for a light beam

$$0 = x^\mu \eta_{\mu\nu} x^\nu = x'^\mu \eta_{\mu\nu} x'^\nu, \tag{238.1}$$

and for every possible inertial frame. The full set of possible Lorentz transformation's $\Lambda^\mu{}_\nu$ are thus those four dimensional matrices which leave $\eta_{\mu\nu}$ unchanged:

$$\Lambda^{T\alpha}_{\mu} \eta_{\alpha\beta} \Lambda^\beta{}_\nu = \eta_{\mu\nu}. \tag{238.2}$$

This includes the Lorentz transformations listed above, plus also ordinary rotations in three dimensional space.

That's all well and good, but how do we do anything useful with this construction? The point is that space and time are not the only things which can be put into a four dimensional structure. The central claim of special relativity is that all physically relevant observables can be expressed as four vectors or quantities that can be derived from four vectors. If we measure the values of these quantities in one inertial frame and coordinate system, then we can compute what values would be measured in a different coordinate system by using the same Lorentz transformations that map the space and time coordinates from one frame to another.

Energy and momentum are another example of a four-vector. Recall from Hamilton's mechanics that there is an analogous relation between energy/momentum and time/space, with energy entering the equations in a similar way to time, and momentum in a similar way to space. This suggests that we should try creating a four-dimensional vector $p_\mu = (E, \vec{\mathbf{p}}c)$, combining energy and momentum. Once again we have had to insert a speed of light into the vector to ensure that each component is measured in the same units; this time it multiplies the momentum rather than the energy. We then know that $p^\mu \eta_{\mu\nu} p^\nu$ is the same in every inertial frame, so let's give it a symbol, say:

$$p^\mu \eta_{\mu\nu} p^\nu = E^2 - \vec{\mathbf{p}} \cdot \vec{\mathbf{p}} = m^2 c^4. \tag{238.3}$$

If we perform a Taylor series expansion for small momentum, we find,

$$E = mc^2 + \frac{1}{2mc^2} p^2 c^2 + \frac{3}{8m^3 c^6} p^4 c^4 + \ldots. \tag{238.4}$$

Once again, we can compare with Newton's mechanics to interpret these terms. The second term is just the Newtonian expression for the kinetic energy, if we identify m with the particle mass. The third term becomes zero as we send the speed of light to infinity, so plays no role in Newton's mechanics. What about that leading order $E = mc^2$? In Newtonian mechanics, you don't ever notice this, because mass doesn't change, no particles are created or destroyed, and you only ever measure energy differences. This is just a constant which you never observe. In other words, this expression reduces to what we would expect in the Newtonian limit.

The first interesting thing is that we get Newton's kinetic energy formula out just by looking for something which is the same in every inertial frame, or Lorentz invariant. The second interesting thing we see is that mass is really just a type of energy, and, conversely, in some circumstances, energy appears like a mass. For example, in a proton or neutron, only a small portion of their mass comes from the masses of the consistent particles; almost all of it is due to the energy binding those particles together. Each particle has a certain amount of internal energy stored within it: and this is what (in non-quantum special relativity)

mass is a measurement of. The next thing we note is that since we can convert energy from one form to another, maybe we can convert energy into mass. But this would mean that we might be able to create and destroy new particles if we have enough energy. That is an interesting idea, but it runs against one of our fundamental mechanistic principles, that matter is indestructible, so let's leave it to one side for the moment.

OK, so we have space, time, energy and momentum sorted out. What about that other component of Newtonian mechanics, force? Let's start with the electric and magnetic forces. There were six basic quantities we needed for Maxwell's equations, the electric charge density ρ, the electric current density $\vec{\mathbf{J}}$, the electric field $\vec{\mathbf{E}}$, the magnetic field $\vec{\mathbf{B}}$, the electric potential ϕ and the magnetic potential $\vec{\mathbf{A}}$. Can we combine these into four vectors? The charge density and current density are the most obvious ones which might be treated in this way. Why? Because current is just charge in motion. So in the rest frame of a particle we have a charge density ρ and no current density. If we then shift to an inertial frame where the particle is moving, we have both a charge and a current which is given by the velocity times the charge. This looks like a Lorentz transformation, so let's try making another four-vector $J^\mu = (\frac{\rho}{\epsilon_0}, \mu_0 \vec{\mathbf{J}})$. Next is the electric and magnetic fields. This is a bit more difficult, since these are described by six numbers, while a four vector has just four numbers. So let's leave them to one side right now. Then we can look at the potentials, and here we have another clear candidate for a four vector, the potential $A_\mu = (\phi, \vec{\mathbf{A}}c)$. Recall that ϕ is an Energy and in Hamiltonian mechanics $\vec{\mathbf{A}}$ is analogous to a momentum, so this is similar to the energy/momentum four-vector. This also actually allows us to deal with the electric and magnetic fields as well, because we note that

$$\partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}, \tag{239.1}$$

with ∂_μ the four vector constructed from the four derivative operators, $\partial_\mu = (\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla})$ and the convention with upper and lower indices being $\eta_{\alpha\mu} F^{\mu\nu} \eta_{\nu\beta} = F_{\alpha\beta}$. This gives,

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z c & -B_y c \\ -E_y & -B_z c & 0 & B_x c \\ -E_z & B_y c & -B_x c & 0 \end{pmatrix}, \tag{239.2}$$

and we can also introduce a similar operator with the electric and magnetic fields swapped:

$$\tilde{F}_{\mu\nu} = \begin{pmatrix} 0 & B_x c & B_y c & B_z c \\ -B_x c & 0 & E_z & -E_y \\ -B_y c & -E_z & 0 & E_x \\ -B_z c & E_y & -E_x & 0 \end{pmatrix}. \tag{239.3}$$

Maxwell's equations then just become

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= J^\nu \\ \partial_\mu \tilde{F}^{\mu\nu} &= 0, \end{aligned} \tag{239.4}$$

I have again omitted the summation over repeated indices. The Maxwell wave equation is

$$\partial^\mu \partial_\mu F_{\alpha\beta} = 0. \tag{239.5}$$

Thus Maxwell's equations, and therefore all of electromagnetism and optics, are already fully consistent with special relativity, since all the key quantities can be expressed in terms of 4 vectors. The Lorentz transformation matrix Λ provides the relationship needed to map the fields and currents from one inertial frame to another. This gives us a new understanding into the relationship between electricity and magnetism – magnetism is just the projection of the electric field into another inertial frame. We can also treat light waves in terms of four vectors – the frequency ω and inverse wavelength $\vec{\mathbf{k}}$ just become another four vector, k_μ, and the location of the peak of the wave $\omega t - \vec{\mathbf{k}} \cdot \vec{\mathbf{x}}$ becomes the Lorentz invariant $k_\mu \eta^{\mu\nu} x_\nu$.

So what about dynamics? Finding equations of motion, and so on. Impact dynamics can be treated in the same way as in Newtonian physics - just instead of conserving energy and momentum, we only need to conserve four momentum. The four momentum is related to the velocity of the particle by $p_\mu = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} m v_\mu$, where $v_\mu = dx_\mu/d\tau$ and the proper time τ is the time as experienced in the particle's rest frame. The movement of particles in an electromagnetic field can be treated in a similar way to Newtonian physics. The particle has electric charge Q. We then find that

$$Q F_{\mu\nu} \frac{dx^\nu}{d\tau} = \frac{dp_\mu}{d\tau} \tag{240.1}$$

gives the Lorentz force law in the limit $\frac{v}{c} \to 0$, and the correct change in energy $\frac{dE}{dt} = \vec{\mathbf{E}} \cdot \vec{\mathbf{v}}$, where the unbolded E is the energy and the $\vec{\mathbf{E}}$ in bold face is the electric field.

The rest of the standard effects of special relativity, which to those prejudiced by Newtonian mechanics and Galilean relativity cause complete confusion, such as time dilation, length contraction, and the speed of light as the universal speed limit follow directly from the equations. For my purposes, such effects are not relevant enough to dwell on. What I want to emphasise are three things.

Firstly, special relativity is essentially just a rephrasing of Newtonian mechanics. The same physical concepts are present, and have the same interpretation. The underlying philosophical principles are, with only two exceptions, the same, but the equations are slightly different because the geometry is slightly different. The principal difference is that it merges space and time together into a four dimensional hyperbolic geometry rather than the three dimensional space plus one time Euclidean geometry of Newtonian mechanics, with everything else ultimately following from that. This is significant in the way that we think about space and time, and it leads us towards a view that past, present and future are all equally real: that it is possible to view them all together as one entity, just as we can view a whole landscape at once, rather than passing by moment by moment in succession. This is, I think, very consistent with the classical view of God existing outside time and able to view all of history at once. Just as we can survey our three spatial dimensions simultaneously, there is no reason why God cannot view the four space time dimensions 'simultaneously' (any difficulty we have in imagining that is a failure of our imagination, which is adapted to a temporal rather than a timeless existence; we can only grasp such things by leaving imagination behind and using the intellect). This, of course, means that God's knowledge of what is to us the future is not dependent on his being able to calculate it, but merely that He observes it along with the present, which also means that God's omniscience does not require the universe to be deterministic.

The second difference is that there is no need for an absolute space or preferred reference frame, i.e. a set of points unchanging in time. Everything is expressed only in inertial frames, using relative space and time. But, with these exceptions, the philosophical world view of classical special relativity is very much that of the mechanism of the early modern era. It adds nothing and takes away nothing from those discussions, except, perhaps, by putting those principles in a more elegant mathematical form.

Secondly, I need to briefly mention causality. Galilean relativity allowed action at a distance; the important thing was that cause had to precede effect, but in principle that cause and effect need not be spatially joined, although there was a lot of discomfort with the idea of action at a distance. The only requirement was that cause had to be a moment earlier in time. In special relativity, this can no longer be correct, because there is now no objective conception of time. But special relativity does say that nothing, including information, can travel faster than the speed of light. This means that we can draw a cone, stretching back from us in time and expanding in space, which is mapped out by the equation $c^2t^2 = x^2 + y^2 + z^2$. So we can imagine light travelling in (at the speed it does in a vacuum) and hitting us at this particular instant of time: every source which could possibly

have emitted that light is on the boundary of this light cone. Everything within that cone can in principle send a signal to us and influence us. Nothing outside it can do so. What this means that any cause of an event had to exist somewhere within that backwards light cone. If the cause happened a moment before the event, then it had to exist at the same location in space. If there is a continuity in time, and the principle of cause and effect holds, then the laws of physics have to be local: no action at a distance is allowed. Equally, there is a light cone spreading out from each event into the future, and, if causality means anything, then any effect of that event can only exist at a space time point within that future light cone.

Thirdly, we should note how symmetry is starting to play a role in physics. Special relativity was derived by constructing a four vector which transforms in such a way as to leave a certain quantity, the proper time $c^2d\tau^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$ invariant. 'Invariant' means that it doesn't change under the transformation; and this is a form of symmetry. There are certain observables which do not depend on which reference frame you are in, and all observers will measure them as the same quantity. Furthermore, having constructed one four vector, we then proceeded by trying to construct everything else in physics as four vectors, or constructed from four vectors and the four dimensional derivative operator, or from the inner product of four vectors, with the implicit assumption that each hanging index transforms in the same way when we change from one inertial frame to another. An invariant is then just something with no vector or matrix indices hanging, with nothing taking a vector product with them. The mathematical form is actually a lot simpler and more beautiful than Newtonian mechanics; and the theory can be largely derived from these mathematical principles alone, although obviously it fits in very well with experiment and we need to turn to experiment to interpret what all the mathematics means. However, we saw aspects of Newton's theory drop out of the mathematical structure; for example Maxwell's equations have to take a similar form to what they do: the connection with the four-vector current (at this level of physics) comes from experiment, but the relationship between the electric and magnetic fields, and the form of the Lorentz force law, are all determined by symmetry. We are no longer in a place where we need to apply induction from experiment to derive physical law. Symmetry and philosophical considerations are just as powerful tools in generating new ideas modern theoretical physics – in many respects more powerful – than concepts induced from experiment, although everything must be tested against experiment.

The most controversial of these points I have raised is with regards to the philosophy of time. There are two general approaches to the philosophy of time, known as the A-theory and the B-theory. In the A-theory, there is an objective meaning to the word "now." There are various different versions of this theory. Firstly, there is presentism, where only the present is real. The past was real but is no longer; possible futures might become real; but only the present can be thought of as actually existing. This is very much in line with our intuitive experience: we experience the present directly, while to access the past we have to rely on memory, and to access the future we rely on reasoned guesswork. Secondly, there is the growing block theory, where the past and present are real, but the future isn't. Thirdly, there could be a shrinking block theory, where the present and future are real, but the past isn't. Fourthly, there is the moving spotlight theory, where in some sense past present and future are all real, but there is something objectively picking out one moment in time as the present. This moves along the time coordinate from one moment to another, but there is still one moment which is objectively picked out as the present. Classical philosophy, and most philosophy before the twentieth century, has adopted the A-theory, and most frequently presentism[13]. It is adopted, in one form or another, by most orthodox Christian theologians, although (in my view) little else in theology actually depends on this theory.

The alternative to the A-theory of time is the B-theory. Here the notion of "now" is only subjective. Everybody, in each location and at each moment of time has their own

[13]For a recent defence of presentism, see (Feser, 2019)

particular conception of the present, but it is internal to themselves, and can differ from the conception held by others in at least some respects (such as which moment of time is the present, or whether two events are simultaneous), but not necessarily all respects. This implies that beings in the past, present and future are all equally part of reality. This view came to prominence with the advent of special relativity, and consistency with the theory of relativity is the main justification for adopting it. The Lorentz symmetry is akin to a rotation, mixing points in space with those in time. Just as we don't view a location in space as objective, because we can rotate coordinate systems (or turn ourselves around) changing which event we assign as being in front of us; so applying a Lorentz transformation (or walking a bit faster) mixes the time and space coordinates, changing which event we assign to the present. If so, then just as there is no objective sense of which place is "here", nor can there be any objective sense of which moment of time is "now", except for beings at precisely the same moment in time and location in space.

Similarly to the A-theory, there are different forms of the B-theory. One, perhaps the most commonly defended and criticised, views time as essentially the same as space (only with that minus sign in the metric). In this view the universe as a whole is just one huge static block, with no real change or movement. But this, I think, goes beyond what is demanded by special relativity, and is inconsistent with other areas of physics such as thermodynamics and quantum physics. Here we find that time is not just like space, but there is a direction and succession in time. If time is the measure of change, and there is a succession in time, with one moment objectively coming after the next, then the universe is not a static block, and there is objective change in the universe. This differs from the moving spotlight view, because the demarcation of one moment in time as the present is a real experience, but subjective (varying from one person to another) rather than objective (the same for every person). It differs from the alternative version of B-theory because the succession and direction of time, and hence the fact of change, is objective rather than subjective. If differs from the growing and shrinking blocks and presentism in that past, present and future are all real, and there is an objective perspective where they have the same ontological status. This version of the B-theory, which is the one I hold, is consistent with both our own experiences of time and with special relativity. In this view, different observers that one event came or will come after another, but they disagree about which event is in the present.

In summary, there are four main questions to answer about the philosophy of time. a) Is the past just as real as the present? b) Is the future just as real as the present? c) Is the notion that one moment of time is singled out as the "present" an objective and fundamental feature of reality, or merely subjective, with the moment of time identified as the present depends only on the individual observer? d) Is there an objective temporal direction and succession? That is, is the notion that one event comes after another an objective feature of the universe (with all possible observers agreeing that one event came first and which event came first), or subjective (with different observers each having their own legitimate and contradictory opinion)? This leaves us with seven logically consistent philosophies of time:

	Real Past?	Real Future?	Objective present?	Objective succession?
Presentism	✗	✗	✓	✓
Growing block	✓	✗	✓	✓
Shrinking block	✗	✓	✓	✓
Moving spotlight	✓	✓	✓	✓
Static spotlight	✓	✓	✓	✗
Dynamic block	✓	✓	✗	✓
Static block	✓	✓	✗	✗

Some people have suggested that the B-theory of time, which is at the very least implied by the geometrical interpretation of special relativity, denies the existence of change and

thus of causality. However, this is only true of the static block theory. The key property of time that implies the existence of change is the notion of objective succession. If temporal succession is merely illusionary, then it makes no sense to say that one event came after another, and hence that one thing caused another thing. However, if temporal succession is an actual and objective feature of the universe, and the direction of time is fixed, then we can distinguish between causes and effects. There is still the possibility of causality. Equally, if we allow for temporal succession, then we can usefully distinguish between something at one moment of time and the same thing at the next moment of time and say that it has changed. If we reject the idea of a temporal direction, then time just becomes like space, and while you can identify difference between one moment of time and another, you can't really call it change because the notion of change implies a hierarchy between the thing that was changed and what it has changed into. So, out of these six theories of time, only the static block version of the B theory of time is inconsistent with the notions of change and causality and implies a Parmenidian changeless universe. The dynamic block version is consistent with the assumption of both Aristotle, Heraclitus and the early modern scientists that change is a real feature of the universe. It is also consistent with the most natural interpretation of special relativity.

While I support the dynamic block philosophy of time, few of the conclusions I will later come to will depend on whether the correct philosophy of time is one of the versions of the A-theory or the dynamic block theory. If, on the other hand, the static block theory (or static spotlight) is correct, then its denial of causality will cause me problems.

7.6. Noether's Theorem

I should mention at this point an important result due to Emile Noether, which is derived from the Lagrangian formulation of classical physics. We extend the Lagrangian formulation to the four dimensional system more convenient to special and general relativity. Thus we have an action

$$S = \int d^4x L(\vec{\mathbf{q}}^\nu, \partial_\mu \vec{\mathbf{q}}^\nu)), \tag{243.1}$$

where $\vec{\mathbf{q}}^\nu$ is some quantity which encodes information about all the particles in the system. It is a vector, and each component of the vector includes information about one individual particle. It has a direction index, and $\vec{\mathbf{q}}^0$ represents information on the time and $\vec{\mathbf{q}}^i$ for $i = 1, 2, 3$ the space coordinates for all the particles. We also include the derivative of $\vec{\mathbf{q}}$ with respect to the location coordinate in some inertial frame. This includes the derivative with respect to time, $\frac{d}{dt}\vec{\mathbf{q}}^i$, which in Newtonian mechanics corresponds to the velocity. This certainly should be in the Lagrangian, but because we want to make everything relativistic, we need to include the other derivatives as well. Note that we have to be a little careful in this: $\vec{\mathbf{q}}$ is not necessarily just a coordinate in our preferred coordinate system, but could represent any vector quantity. Following Hamilton, we write the canonical momentum density (i.e. momentum per unit volume, since we introduced an integral over space into the action this is the Lagrangian density rather than the Lagrangian) as $\vec{\mathbf{p}}^i = \frac{\partial L}{\partial \frac{d}{dt}\vec{\mathbf{q}}^i}$, and to be consistent with the four vector notation, we write a canonical energy density as $E = \frac{\partial L}{\partial \frac{d}{dt}\vec{\mathbf{q}}^0}$.

Suppose that we change $\vec{\mathbf{q}}$ in such a way that the action remains unchanged. Recall that any transformation that leaves the action unchanged does not affect the underlying physics. We write

$$\vec{\mathbf{q}}^\nu \to \vec{\mathbf{q}}^\nu + \epsilon w^\nu, \tag{243.2}$$

where w parametrises the change and ϵ is some small constant shift. This means that the change in the Lagrangian is

$$\epsilon w^\nu \cdot \left(\frac{\partial L}{\partial \vec{\mathbf{q}}^\nu} - \partial_\mu \frac{\partial L}{\partial \partial_\mu \vec{\mathbf{q}}^\nu} \right) + \epsilon \partial_\mu \left(\omega_\nu \cdot \frac{\partial L}{\partial (\partial_\mu \vec{\mathbf{q}}_\nu)} \right). \tag{243.3}$$

The first of these terms is zero according to the equations of motion of the system – it is just the relativistic version of the Euler-Lagrange equation. The action is unchanged, which means that the second term must also be zero, as ϵ is arbitrary.

The quantity $\frac{\partial L}{\partial(\partial_\mu \vec{q}_\nu)}$ is known as the stress energy tensor, and is given the symbol $T^{\mu\nu}$. It's importance can be seen by looking at the components of the matrix where μ represents time: we then have T^{0i} is $\frac{\partial L}{\partial(\frac{d}{dt}\vec{q}_i)}$ which is just the momentum density, and T^{00} is correspondingly the energy density. The matrix is actually symmetric, which means that T^{i0} is also the momentum density. the space/space components T^{ij} are known as the pressure, stress or shear depending on the component and circumstances.[14]

This means that, introducing a integral over all space,

$$\frac{d}{dt}\int d^3x w^\nu \cdot \frac{\partial L}{\partial \frac{\partial}{\partial t}\vec{q}^\nu} = \int d^3x \partial_i \left(w^\nu \cdot \frac{\partial L}{\partial \frac{\partial}{\partial x^i}\vec{q}^\nu} \right). \tag{244.1}$$

The right hand side of this equation, by Stokes' theorem, just gives the surface integral over the boundary of all space. But space doesn't have a boundary, so this is just zero. What this means is that,

$$\frac{d}{dt}\int d^3x w^\nu \cdot \frac{\partial L}{\partial \frac{\partial}{\partial t}\vec{q}^\nu} = 0, \tag{244.2}$$

or,

$$\int d^3x w^\nu \cdot \frac{\partial L}{\partial \frac{\partial}{\partial t}\vec{q}^\nu} \tag{244.3}$$

does not change in time (remember that we have previously specified some inertial coordinate system, and that the differential function $\frac{d}{dt}$ is a measure of change). Let's take an example. Suppose that w^ν represents a shift in time. We just move everything in the universe forward one second. This can't make any difference to the physics, since physics only depends on relative duration, and thus the action must be unchanged and this equation applies. What this means is that we are looking at the $\nu = 0$ component of this equation, and from Hamilton's mechanics we know that $\frac{\partial L}{\partial \frac{\partial}{\partial t}\vec{q}^0}$ is the canonical energy density. So what this tells us is that the total energy in the universe does not change in time. w does not change in space, $w^0 \frac{\partial L}{\partial \frac{\partial}{\partial t}\vec{q}^0}$ gives the sum of the energy of all particles at one particular point in space, and $\int d^3x$ means a sum over all points in space. Thus this whole equation represents the total energy in the universe at one particular moment in time (in any reference frame), and what we have learnt is that this does not change in time. Thus we have just proved based on only the symmetry with respect to constant translations of the time coordinate and the principle of least action that energy is conserved.

Similarly, we could make $\vec{w}$ a shift of every particle in the universe one centimetre to the right. This means that we are looking at a spatial component of ν, $\frac{\partial L}{\partial \frac{\partial}{\partial t}\vec{q}^\nu}$ represents a component of the momentum, and we have just proved momentum conservation. If we considered a rotation, we would show conservation of angular momentum. I mentioned gauge transformations in electromagnetism (equation (220.1)). This is effectively another symmetry of the action, and implies the conservation of electric charge.

In other words, the conservation laws assumed by Newton in order to make sense of empirical observations (and encoded in his third law), can actually be derived abstractly from a very different principle. Newton got his laws from experiment; now we are almost in a position where we can get them from philosophy. We have made two assumptions: firstly the

[14]Note that in general relativity, an alternative definition of the Stress Energy tensor is usually used, the differential of the Lagrangian with respect to the metric tensor. The two definitions are equivalent, as long as we are careful with gauge transformations while we translate. For every translation of an electromagnetic field, we also need to simultaneously perform a gauge transformation. If we get this right, then one obtains the same stress energy tensor by both methods.

principle of least action, and secondly that certain symmetry transformations don't affect physics. These symmetry transformations are basic: Lorentz invariance, translations, and so on. Ultimately they are the sort of thing that can be derived from philosophical principles. For example, one might argue that God is timeless and outside space. It shouldn't matter to God whether we give an event a time coordinate of one second or two seconds, as long as we shift the time coordinates of everything else in the universe forward by a second as well – only differences in time are observable, not absolute coordinates; God doesn't provide an absolute coordinate system. If the 'laws of physics' describe how God upholds the universe, then these too should be unchanged under translation, rotation, etc., since God is not affected by such trivialities and thus His upholding of the universe won't be either. But these symmetries then give us the conservation laws, and the conservation laws are the foundation of classical physics (providing us with Newton's second and third laws). Of course, there is still the principle of least action. Why should physics obey a least action principle? Can we derive that from some deeper principles? If we can, then philosophy alone might be enough to at least partially constrain the laws of physics: we would not need induction from experiment at all, if only we were smart enough. We are, however, not yet in a position to attempt this. But I will get to it in the appropriate place.

7.7. General Relativity

My discussion of special relativity left out one important aspect of Newtonian mechanics, namely gravity. It might be thought that this is not such a big problem. After all, electrostatics has a very similar mathematical form to Newton's law of gravity, so maybe gravity can be treated in precisely the same way. All we have to do is introduce a gravitational analogue of the magnetic field, which we can perhaps suppose was not observed because things don't travel fast enough so any effect it had was beneath the experimental precision, a gravitational four vector potential, and plug in an exact equivalent of Maxwell's equations. And people thought of this, tried it, and discovered that it did not work out.

Einstein himself was never really tempted to go down this route, or if he was tempted it was only very briefly. He was troubled by one interesting point about the gravitational force: it was proportional to the mass of the particle. If gravity could be formulated in the same way as electromagnetism, then there would be no obvious reason for it to be so. It would just have to be some huge coincidence, and coincidences on that scale lead us to seek alternative explanations. Equally, that gravity is always attractive, while the electromagnetic force can be either attractive or repulsive, again suggests that there is a fundamental difference between gravity and electromagnetism. Einstein instead noted something curious: that it was impossible to distinguish between gravitational force and acceleration. If you are in elevator (or anything else) in free-fall, you don't experience any gravity. You are falling at the same rate as your surroundings. Equally, if you were in deep space and your spacecraft was accelerating, the 'floor' of the spacecraft would be pushing up against you, and you wouldn't be able to distinguish whether this was an accelerating spacecraft pushing up or the force of gravity pushing you down. You wouldn't be able to tell whether or not you were on the spacecraft or on a planet.

It was reasoning such as this which got Einstein thinking about accelerating frames of reference (for an introduction to general relativity see (Einstein, 1920; Wald, 1984; Sklar, 1974; d'Inverno, 1992)). What he determined to do was to expand his principle of invariance from inertial frames to any general accelerating frame and coordinate system, with the hope that a theory of gravity would drop out of this at the same time. Fortunately for Einstein, a lot of the mathematics he needed had already been done in the nineteenth century. There were still philosophical roadblocks: his mind was tuned to Newton's idea of absolute space and time, and the realisation that there was not even something special about inertial frames took some time for him to get his head around. But his end result is a piece of mathematics of the greatest beauty. Classical general relativity is not the full story; it needs to be

incorporated into quantum mechanics, and there things start getting a lot messier. But I almost wish that it was the full story, because (like quantum field theory) the mathematics is just so elegant while all the candidate quantum gravity theories I know of add layers of mathematical complexity and ugliness.

I mentioned that there were several forms of non-Euclidean geometry discovered in the nineteenth century. We have already mentioned the hyperbolic geometry of Gauss and others as the foundation of special relativity. Now we need to turn to Riemann and his differential geometry.

If we are to equate acceleration with gravity, then we must extend special relativity to general coordinate systems: we replace the Minkowskian coordinates x_μ with $\tilde{x}_\mu(x)$, where $\tilde{x}_\mu$ may be any valid function of x_μ, not just the linear transformations considered in special relativity. $\tilde{x}_\mu$ is constrained by two conditions: firstly that the mapping from x_μ to $\tilde{x}_\mu$ is reversible, i.e. it is one-to-one (every value of x_μ corresponds to precisely one value of $\tilde{x}_\mu$, and vice-versa), and that the mapping preserves the topology of space time (for example, it should not introduce any new discontinuities into the description of space and time). It need not preserve the metrical structure, but the determinant of the transformation matrix $\frac{\partial x_\mu}{\partial \tilde{x}_\mu}$ must not be zero or infinite at any point in space or time. We must also replace Newton's first law with something that extends to any coordinate system; not just inertial frames. What we will find is that (classical) gravity is very different from how Newton envisaged it: it is not a force in the same way that electromagnetism is a force, but a distortion of space and time. Newton said that free particles travel in straight lines, the shortest distance between two points. It turns out that planets in orbit around a star obey the same principle: it is not the path of the planet which is curved, but space which is curved. The planets continue to travel in 'straight' lines[15].

This seems rather counter-intuitive and confusing. We can consider an ant walking along the surface of a sphere. He cannot move through the interior of the sphere, but is constrained on the surface. The directions he can walk in are determined by the tangent vectors of the sphere, so at each point in the sphere there is a North/South direction and an orthogonal East/West direction, although we may choose any arbitrary point as the North pole. The shortest distances between two points on the sphere would involve walking along the same tangent vector relative to the surface. So, for example, if you are on the equator the shortest distance to reach another point on the equator is to keep walking East (or West). The equator corresponds to one of the geodesics of the sphere. If you continue to walk East, you will, of course, eventually come back to where you started. Thus although you continue to walk in the same direction relative to the sphere surface you nonetheless eventually return to your starting point. We look at the sphere surface from the viewpoint of flat space time; a curved surface is very different. A curved surface basically means that the coordinates for the surface of the sphere are not linear functions of the equation which describes the surface of the sphere (e.g. $x^2 + y^2 + z^2 = 1$)[16]. However, there is another perspective, which is entirely equivalent though takes some getting used to, that the sphere forms a 'flat' surface in curved space time. This means that the points on the surface of the sphere can be described by a linear function of the coordinates in our spherical coordinate system (the equation for the surface is $r = 1$ where r is the radius). However, there must be curvature somewhere, and it is contained within the definition of the coordinate system, which can be characterised by the metric. In flat-space time[17], the metric is a linear function of the coordinates; in curved space time the metric is a non-linear function of the coordinates.

---

[15]Curiously, this is similar in effect to Buridan's extension of the principle of inertia to allow constant rotational motion; he explained the orbits of the planets as being their inertial motion. The explanation in general relativity is not so different; the planets, like everything else, continue along geodesics unless acted upon by an external (non-gravitational) force.

[16]In practice this is a poor definition of curvature, but for the present discussion it is good enough.

[17]Again, I am being a little imprecise here, and will be more definite later.

Another picture of general relativity is that space-time is like a (frictionless) rubber sheet that contains small dimples (but no bumps). We naturally think of it just as a flat or planar sheet, so if you roll a ball across the surface of the sheet it will continue rolling in the same direction. The tangent vectors of the sheet are along the surface of the plane. However, if there are small dimples in the surface, it will change its direction when encountering one of these depressions, since the tangent vectors of the surface no longer point along the plane. It is even possible in a frictionless system (which never really exits in practice) that the marble may proceed along a path trapped in one of these depressions and continue circling in a fixed orbit; this is equivalent to the orbit of a planet around a sun.

The second of the principles upon which general relativity is based is Mach's principle, which states that in the absence of matter there should be no meaning to the concepts of rotation and acceleration. This effectively means that physics (until we introduce particular configurations of matter) ought to be invariant under generalised coordinate transformations. General relativity states that the intrinsic, observer independent properties of space time are described by a metric, as in special relativity, but that this metric is not flat. The effects of the curvature of matter are what in Newtonian mechanics are accounted for by the gravitational field. Particles 'naturally' travel along geodesics of the metric (defined below); and the curvature of the metric is related to the stress-energy tensor.

General relativity does not mean that all reference frames are equal: particles in some reference frames experience gravitational forces, while in other frames they don't. Even the indistinguishably of inertial frames disappears in general relativity except at individual points in space time. Local Lorentz transformations remain a symmetry of the Lagrangian, global Lorentz transformations don't. This means that there is a preferred reference frame at the cosmic scale, defined by the cosmic microwave background, the residue from the singularity that marks the beginning of the universe.

7.7.1. Metric Tensors and Curved Space Time

In special relativity, the path of light is given by the equation

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2, \tag{247.1}$$

where s parametrises the path of the light, and dx indicates a small change in the location of the beam. This may be written as

$$ds^2 = dx^\mu g_{\mu\nu} dx^\nu, \tag{247.2}$$

where $g_{\mu\nu}$ is the metric tensor, $g_{\mu\nu} = \text{diag}(1, -1, -1, -1) \equiv \eta_{\mu\nu}$. It is conventional to write $x_\mu = g_{\mu\nu} x^\nu$, since this expression appears frequently. One therefore contracts a raised index with a lowered index to get a scalar. We have already seen that Lorentz transformations, equivalent to a change in velocity, are changes which leave the metric invariant.

We now, however, wish to consider changes to the coordinate system which do not leave the metric invariant. We have the original, Minkowskian, coordinates x_μ. We now switch to a coordinate system $\tilde{x}_\mu$, but demand that ds^2 is invariant. We find,

$$ds^2 = d\tilde{x}^a \frac{dx^\mu}{d\tilde{x}^a} \eta_{\mu\nu} \frac{dx^\nu}{d\tilde{x}^b} d\tilde{x}^b, \tag{247.3}$$

which in effect gives us a new metric

$$g_{ab} = \frac{dx^\mu}{d\tilde{x}^a} \eta_{\mu\nu} \frac{dx^\nu}{d\tilde{x}^b}. \tag{247.4}$$

These new coordinates are a function of the old coordinate, but there is no reason why they have to be a linear function (i.e. a corresponding to a translation, change in velocity or rotation). For example, the new coordinates could correspond to a reference frame which is accelerating with respect to the original one. We can think of two observers, one on a bench and one on a train; each devises a coordinate system with themselves at the centre, only this time the train is accelerating according to the observer on the bench; while the observer on

the bench seems to be accelerating according to the observer on the train. So are the two cases mirror images of each other? Not quite, because unlike linear transformations (e.g. two observers travelling at a constant relative velocity), each observer will measure a different metric tensor that governs the path of light waves. For example, if the observer on the bench calculates that the laws of light as he measures them obey the Minkowski tensor, then the train observer will measure a different metric tensor. There isn't any physical difference between the two observations: each observer can take their own measurements, apply the rules of general relativity, and accurately compute what the other measures. They measure the same thing; it just seems to be different because of their different (arbitrary) choice of coordinate system when creating an algebraic representation of physical space time.

Switching to this coordinate system will also modify the differential operator. In particular, when we convert between two coordinate systems, we need to modify the differential operator to account for this difference. Since the laws of physics are written in the form of differential equations, this means that the symbols written down to describe physics will depend on the reference frame. What we are interested in is how to relate the equations of one reference frame with those of another. We use Minkowski space time from special relativity as the first reference frame – partly because it is a very good approximation of the world we live in, but mostly because we know how to formulate physics in Minkowski space time. The differential operator in this frame is the usual ∂_μ operator that dates back to Newton and Leibniz and the early days of calculus. The second reference frame is that of an arbitrary coordinate system, and we call the natural differential operator in this frame ∇_μ, and define it by two properties: (1) it is the appropriate transformation of ∂_μ into the new coordinate system; and (2) the differential of the metric tensor in the new coordinate system is always zero, $\nabla_\mu g^{\mu\nu} = 0$. This second condition is satisfied in the flat space time cases we know about.

The two differential operators are related by something known as the Christoffel symbol denoted by the symbol Γ. When acting on a four-vector ω^b (this might represent energy/momentum, or electric/magnetic potential, or charge/current density), the appropriate equation to map between the two is

$$\nabla_e \omega^b = \partial_e \omega^b + \Gamma^b_{de} \omega^d$$

where the Christoffel symbol is defined by

$$\begin{aligned}\Gamma^b_{de} =& \frac{\partial \tilde{x}^b}{\partial x^\rho} \frac{\partial^2 x^\rho}{\partial \tilde{x}^d \partial \tilde{x}^e} \\ =& \frac{1}{2} g^{ba} \left(\tilde{\partial}_e g_{ad} - \tilde{\partial}_a g_{de} + \tilde{\partial}_d g_{ae} \right). \end{aligned} \tag{248.1}$$

For every linear transformation of the coordinates (translation, rotation, change in velocity) the Christoffel symbol is zero, and the form of the equations that describe physics are unchanged. But in an accelerating reference frame it is not zero. What significance does this have? Let us consider Newton's first law, which states that in an inertial frame (i.e. governed by the Minkowski metric) a body continues with constant motion unless acted upon by an external force. If we make the four vector ω^b used above the energy/momentum four vector and consider the time component of the derivative operator, Newton's first law converted to equations states that in the absence of a force $\partial_t p^b = 0$. p^b represents energy and momentum, so this equation states that the time derivative of energy is zero - the energy of the particle is constant – and the time derivative of the momentum is zero – the momentum of the particle is constant, and thus the velocity doesn't change. This is just what we mean by inertial motion. But if we then switch to an accelerating reference frame, we find that the equivalent equation is

$$\nabla_t p^b = \Gamma^b_{dt} p^d \tag{248.2}$$

In other words, observers travelling in this reference frame don't naturally travel with constant velocity (as measured in the $\tilde{\mu}$ coordinate system): they will seem to be accelerating. Of course, this apparent acceleration is just a consequence of the fact that they have made life harder for themselves by choosing a more complicated way to convert from physical space time to a geometrical representation to an algebraic representation. If they had just chosen a straight forward inertial frame with a Minkowski metric, they would have found things so much easier. Our premise is that there is no external force, so from this perspective we could say that Newton's Laws are violated in an accelerating reference frame.

But there is another way of reading equation (248.2), and that is to say that Newton's laws are still valid in any inertial frame, but that the measured forces depend on the coordinate system. In other words, we treat $\Gamma^b_{dt}p^d$ as a force acting on our body. Indeed, it has the properties of a gravitational force, in that it is proportional to the body's mass (technically it is proportional to the Energy, but the Energy is dominated by the mc^2 mass contribution). This term is, of course, only the artefact of switching from an inertial to an accelerating coordinate system. So this suggests that rather than saying that Newton's laws of motion are only valid in certain coordinate systems (as Newton and Galileo did), we could say that Newton's laws (which from a mechanistic perspective encode the Mechanical Laws of Physics) are valid in every coordinate system, but our perception of gravity depends on the coordinate system we use. This is why when we are in a car undergoing linear acceleration, we feel that acceleration; when on a roundabout, we feel the centrifugal and Coriolis forces: these are all essentially just extra gravitational forces caused by our point of view being trapped in a non-inertial reference frame. These all appear in the appropriate Christoffel symbol for the particular accelerating reference frame which describes our own life. So we now have two classes of gravitational force: the one we are used to keeping us trapped on the earth, and the earth in orbit around the sun; and the second class (in Newtonian mechanics called a 'fictitious force') which might arise from making a strange choice of coordinate system.

But of course, these two classes of gravitational force are not so different after all; we can choose a reference frame such that the gravitational force from the Earth and sun is exactly cancelled out by the fictitious forces from the Christoffel symbol. In this reference frame, there would be no gravity. The earth would, of course, still be in orbit around the sun, but in this crazy coordinate system it would experience no force: what to us would seem to be circular motion would to somebody using those coordinates be inertial motion of constant velocity. This is like someone in a falling lift or a spacecraft in orbit around the Earth: they feel no gravitational force because they are in an accelerating reference frame which precisely cancels out the effect of gravity.

Or here is another crazy thought: we have been assuming throughout this discussion that (once we have compensated for the minor effects of the rotation of the earth and any acceleration we are in relative to the surface of the earth) we are naturally in an inertial frame – the Minkowski metric applies to us. What if that is not true; what if we are in curved space time, and thus we consequently ought to be adding a Christoffel symbol to all the differential equations that describe physical law – and we do by adding into our equations the standard gravitational force. Gravity is not an external force as Newton supposed, just a 'fictitious force' due to a Christoffel symbol. This might work; looking at equation (248.2), if the Christoffel symbol is such that Γ^b_{xt}, Γ^b_{yt} and Γ^b_{zt} are insignificant in comparison to Γ^b_{tt}; which means that the fictitious force would be dominated by the time component of the four vector denoted p^d, which represents the energy of the particle, which is dominated by its rest mass. Thus the force of Gravity would (baring an incredibly accurate experiment) seem to be directly proportional to the particle's mass, which is what Galileo and Newton suggested. The orbits of the planets is not because of a force acting on those planets, but because of an intrinsic curvature of space-time.

If we are going to continue this line of thought, we need a few clarifications. Firstly, we need to define what a straight line is in a general coordinate system (so we can properly define inertial motion in any coordinate system). Secondly, we need to establish what we mean by curved space time. Thirdly, Gravity is based on the idea that masses mutually attract each other (or each mass generates a gravitational field in Newton/Maxwell's field representation of the forces), which in this description means that masses generate a curvature of space time. We need an equation which describes the curvature generated by matter.

The definitions of a straight line and curvature were worked out in the nineteenth century by the mathematicians. Suppose that t^a is a tangent vector representing the direction of a straight line. Then, in a general coordinate system, this vector satisfies the equation $t^a \nabla_a t^b = 0$. Curvature also has a precise mathematical definition. The standard (Newton/Leibniz) derivative operator satisfies the property

$$(\partial_a \partial_b - \partial_b \partial_a)\omega_c = 0,$$

which means that it doesn't matter if you consider how something moves in the x-coordinate first and then how it moves in the y-coordinate second, or movement in y first and in x second; you will get the same result at the end. This is satisfied in both Euclidean (i.e. Newton's mechanics) and hyperbolic (i.e. special relativity or Minkowskian mechanics) geometry. By flat space-time, we mean a geometry where this condition is satisfied by the natural derivative operator of that coordinate system. But in a general coordinate system, equation (7.7.1) is not satisfied. Instead we find,

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)\omega_c = R_{abc}{}^d \omega_d,$$

where $R_{abc}{}^d$ is known as the Riemann curvature tensor. It can be expressed in terms of Christoffel symbols

$$R_{abc}{}^d = \partial_a \Gamma^d_{bc} - \partial_b \Gamma^d_{ac} + \Gamma^\sigma_{ac}\Gamma^d_{\sigma b} - \Gamma^\sigma_{bc}\Gamma^d_{\sigma a}, \tag{250.1}$$

and therefore in terms of the metric tensor. If the Christoffel symbol is zero, then the curvature tensor will also be zero; if the Christoffel symbol is not zero, then curvature tensor might be zero, but most of the time it won't be. This tensor defines what we mean by the curvature of space-time. A further, slightly simpler, measurement of the curvature of space-time is the Ricci tensor

$$R_{ac} = \sum_b R_{abc}{}^b, \tag{250.2}$$

and also of interest is the Ricci scalar which will also only be non-zero if space time is curved,

$$R = g^{ab} R_{ab}. \tag{250.3}$$

The curvature of space time is precisely and mathematically defined by the Riemann and Ricci tensors. If these are zero, there is no curvature. In an accelerating reference frame, there is curvature.

All of this had been worked out by the nineteenth century mathematicians. Einstein's contribution was to realise that the mathematics was related to gravity, and to write down the equation that describes how space time is curved by matter. This is known as Einstein's field equation, and here it is:

$$R_{ab} - \frac{1}{2} g_{ab} R + \Lambda g_{ab} = \frac{8\pi G}{c^4} T_{ab}. \tag{250.4}$$

R_{ab} and R are the Ricci curvature tensors, g_{ab} is the metric tensor, Λ is known as the cosmological constant (which Einstein described as his 'greatest mistake' when he modified his theory to put it in to allow a static universe, then took it out again when it became clear that the universe was not static, but later cosmologists put it back in again when it was measured to have a very small but non-zero value), T_{ab} is the stress-energy tensor: the Noether current associated with translations in space or time, G is the same constant that

appears in Newton's theory of gravity, and c is the speed of light. The stress energy tensor, as the name implies, contains as one of its components (the time-time component) the energy density of matter. Mass is a form of energy, and it happens that if we accept a heliocentric coordinate system the stress energy tensor of the sun is (to a good approximation) zero except for the time-time component, which is (to another good approximation) proportional to its mass. So what this equation states is that mass curves space time, which in turn creates a Christoffel symbol for every other piece of matter in the universe, which, as we have already seen, is mathematically indistinguishable from a gravitational force.

If the space-time curvature is sufficiently small, Einstein's theory reduces to the equations that describe Newton's law of universal gravitation. If we take the limit that the speed of light becomes infinite in comparison to the observed velocities of the planets and stars in our coordinate system, write the metric of space time as the Minkowski metric plus a small correction, and neglect terms proportional to the square or higher powers of that correction, we get something mathematically indistinguishable from Newton's theory of gravity. Of course, if we don't neglect those square powers, we get a slightly different expectation -- and those differences are observable, and side with Einstein rather than Newton. Thus Einstein's theory of gravity both explains why Newton's theory works so well, and improves on it. The general theory of relativity has been tested experimentally to an incredible precision, and passed every test. Not only that, it is mathematically a quite beautiful theory: almost all of it just follows from the principles of geometry plus the principle of inertia, modified to apply to general coordinate systems. It removes Newton's 'spooky' action at a distance. Indeed, I would say that if we assume mechanism and that space, time, and locomotion can be treated geometrically, then general relativity has to be part of the laws of physics. It represents the culmination of the philosophical ideas that inspired Galileo and Descartes; and only awaited the underlying mathematics to be discovered in the late nineteenth century, and a physicist of sufficient wisdom and ability to apply them.

Of course, there are still unanswered questions: why do the constants such as Λ and G take the values that they do[18]? Why does matter curve space time as described by Einstein's field equation (if indeed it does: there could in principle be higher order corrections coming in at levels below our current experimental precision); indeed why does it curve space-time at all? Thus General relativity is not the full answer; although many of its principles arise from geometrical considerations and thus could hardly be different, there are parts of it which are incorporated as a response to the needs of experiment, and parts due to the demands of a mechanistic metaphysics. There is certainly at least one layer to reality more fundamental than classical general relativity, though what that is is still unknown.

General relativity is the culmination of the ideas that began with Bradwardine and the Merton calculators almost seven hundred years earlier: that physics, both motion, gravity and solar system dynamics, is best treated geometrically. Of course, it required numerous advances of mathematics before this dream could be realized, the most important being the discovery of calculus. Indeed, in many cases, the mathematics was developed in response to the needs of physics. On the face of it, it does not stand against the mechanistic world view, or add anything to it, but rather represents its perfection.

Of course, there are still a number of assumptions behind classical general relativity, and it is worth listing a few of them

(1) That matter is made from indestructible particles which are bound to obey laws of motion external to them. (The mechanistic principle)
(2) That there is general coordinate invariance; namely that there is no single absolute frame of reference, including accelerating frames, but physics in one frame can be mapped to another using the principles of Riemann geometry. (The principle of objectivity, no preferred observer.)

[18]At least, in comparison to other physical scales: since they are dimensional quantities the question only makes sense as a comparison.

(3) Space and time can be mapped to a four dimensional geometry with a Minkowskian metric signature (i.e. the sign of the term in the metric corresponding to the time-time component is opposite to the sign of the space/space components; the actual value of the metric will not be Minkowskian but represent some curved geometry). (The geometrisation of physics)
(4) The force of gravity is equivalent to an accelerating reference frame; and one can choose a reference frame such that the force of gravity is absent. (The principle of equivalence).
(5) That matter proceeds along geodesics unless acted on by an external force. (Newton's first two laws, updated).
(6) That matter interacts with the curvature according to Einstein's field equation.

General relativity does, however, have several consequences. There are several known solutions of Einstein's field equation. Of most interest is the small gravitational field solution, as already discussed, because this solved discrepancies between Newton's theory and observations of the orbit of Mercury, and provided an early experimental confirmation of general relativity. Other experimental tests of general relativity have been performed, and so far it has passed them all to an exceptional accuracy. One well known solution, the Schwartzchild solution, predicts the existence of black holes. The metric most often used in cosmology, the Robertson-Walker metric, indicates that the universe cannot be eternal in time or static, but must have a beginning; and a celebrated theorem by Penrose and Hawking (Hawking and Penrose, 1970) confirms that classical gravity is inconsistent with a universe stretching eternally back into the past. This is perhaps the consequence of general relativity that most caught people by surprise.

It was long held - by Aristotelians as well as the mechanists that one could not prove that the universe had a beginning, but that it was fully consistent with it having existed for ever in a steady state. Christians were told that the beginning of the universe was something we had to accept on the basis of revelation, as it could not be proved by reason. With a series of discoveries through the twentieth century, starting with Hubble's observation that the universe was expanding. This means that the distances between galaxies is increasing due to changes in the space time metric – the picture one should have is to make marks on a rubber sheet to indicate galaxies, and then to move them apart by stretching the rubber sheet. The marks (which represent galaxies) remain in the same place relative to the rubber sheet (which represents space-time), but from our perspective the distances between them seem to be increasing. It is not galaxies that move in space, but that space itself expands[19]. The more recent observation of the cosmic microwave background is seen as the definitive experimental confirmation of the big bang picture of the universe's origins.

Thus one thing that Aquinas and the medieval scholastics got wrong was the idea that it is impossible to demonstrate that the universe had a beginning through secular reasoning; instead it follows as a consequence of the geometrical principles underlying general relativity. It has been proved that if classical general relativity is true, the universe had a beginning. Of course, whether classical general relativity is true is another matter. The problem is that if we extrapolate back in time to the earliest moments of the universe, we know that

[19]In geometry, the metric provides the scale we use to measure the distance between two points. This scale is not independent of time, as it is in the hyperbolic geometry of special relativity, but depends on time. By the universe expanding, we mean that this measure of the distance between two points decreases as time progresses. This is difficult to visualise, but one way of picturing this (which, admittedly, like every analogy has its problems) is to say that it is not that the galaxies move with respect to the underlying space; rather the natural rulers used to measure their distance shrink in every frame of reference consistent with Einstein's field equation. I mean by a *natural ruler* that it is implied by the original choice we made of how to map from physical space to geometrical space. That choice of mapping is not completely arbitrary, but has to be consistent with Einstein's field equation. We cannot change the nature of that mapping from one time to another; we have to be consistent. That mapping provides a way of comparing standard length scales from one time to another or one point in space to another, and from that comparison we find that those length scales decrease in size as we move to later time periods.

classical gravity is going to break down and that quantum gravity takes over. We do not know what that theory is. It is likely that the conclusion that the universe has a beginning will remain (it still applies in various candidate theories of quantum gravity) but until the fundamental theory is established, there is always an element of doubt (whether one is looking for an eternal universe or a finite universe).

7.7.2. Lagrangian Formulation

Newtonian Mechanics and special relativity are most clearly expressed in terms of a Lagrangian. The equations of motion are determined as the trajectories which minimise a scalar action. Can Einstein's field equation be expressed in terms of a Lagrangian? It can, quite easily. The most common action used for General Relativity is the Einstein-Hilbert action, the integral over space and time of the Lagrangian:

$$\mathcal{L} = \sqrt{-g}R + \sqrt{-g}\mathcal{L}_M, \tag{253.1}$$

where $\mathcal{L}_M$ is the Lagrangian of the matter (expressed in some coordinate system invariant form) and g is the determinant of the metric tensor (which is negative). The $\sqrt{-g}$ is to ensure that the integral over space time, now $\int d^4x\sqrt{-g}$, is invariant under transformations of the coordinate system. We add to this term the correct Lagrangian for the matter (for example Maxwell's Lagrangian for electromagnetic fields and the Kinetic energy term for matter fields). This Lagrangian will then fully describe both classical gravity and classical electromagnetism. Minimising the action with respect to the metric gives Einstein's field equation (with the stress energy tensor arising from the change of the matter Lagrangian when the metric is changed slightly). Minimising the action with respect to the coordinates that describe the matter and the fields gives the geodesics and appropriate electromagnetic force terms for the particles.

Of course, this formulation carries an important (and to me non-intuitive) implication: the metric of space time is treated as a dynamical variable of the system. This is the source of the difficulties when trying to create a quantum theory of gravitation. The natural impulse, given this Lagrangian, is to try to turn the metric itself into a quantum field, but this is both unnatural (since the metric represents the representation of a property of space-time rather than a physical object) and runs into mathematical inconsistencies (the theory produces infinite results. There is nothing unusual with infinities in quantum field theory, but many of them can be tamed by careful tuning of the various parameters of the system. In the case of a naive quantum gravity, this approach which works so well in other contexts isn't possible).

CHAPTER 8

The Philosophy of Classical physics

8.1. Positivism

Positivism can be defined as the belief that it is possible to acquire truths about things without presupposing any theory of its ultimate nature, and that it is possible to isolate and study one part of the phenomena without needing to understand the whole. This is at the centre of Newton's method, and it has to a large extent been carried out throughout classical physics. Indeed, it would be impossible to pursue the experimental method without assuming this. Suppose I want to perform an experiment: the first thing I do is seal off my equipment so that the only influences on the result are those things that I control. This is essential for the experiment to be replicable. If (for example) the amount of traffic outside affects the experiment, then I will get different results depending on whether or not it is the rush hour. This will help me study traffic flows, but not much else. But if I am going to isolate my experiment from the rest of the universe, I have to presuppose that it is possible to do so. It may be thought that the success of experimental science shows that it is possible to do so, but this is begging the question. At best it has been shown that it is possible to do so for the experiments conducted so far; I am not even sure that we have gone that far, since the possibility to isolate experiments is a presumption of the theory, and if we find it to hold that might mean that we are reading it into the experiment (for example, in how we set up the experiment) rather than out of it.

The second aspect of positivism is the idea that it is possible to study the parts of a complex substance individually, without thinking of what changes when we put those parts together. In many fields, studying the parts in isolation is useful; in others not so much. This ties into the mechanistic principle that a being can be reduced to the sum of its parts.

A further use of this principle is in the separation of scales. For example, if we are interested in large scale properties of substances with a lot of particles - variations of the sort of distance scales we are used to – we can do so very well using thermodynamics, which derives relations concerning the average properties of (for example) gases. However, when we move to a smaller length scale, then we need another theory, statistical mechanics. You can take statistical mechanics, the microscopic theory, and derive from it the conclusions of thermodynamics, the macroscopic theory. Everything in thermodynamics is contained within statistical mechanics; and in this sense statistical mechanics is the more fundamental theory. However, in order of discovery, thermodynamics came first; we would not have been able to think of the lower distance theory without first understanding the large scale theory. That we can think of the large scale theory in isolation of the microscopic details is thus both useful, because it means that we can make good progress one slice of reality at a time, and irritating, because it makes finding experimental evidence of the more fundamental theory harder.

It is sometimes thought that positivism does away with metaphysics. After all, if we can acquire information without presupposing the grounding of physics, then it can't really matter what lies beneath physics. And if physics (together with those other sciences built on physics) explains everything we can observe, then there isn't much hope for speculations about metaphysics to be proved true or false. This idea, however, is dangerous and should be avoided. It is dangerous because it leads to sloppy thinking. The idea that it is possible

to acquire truths without recourse to theories of ultimate natures is in itself a theory about ultimate natures. For this will only hold for certain ultimate theories of nature; indeed even to precisely define the phrase 'ultimate theory', if only to deny its existence, we need metaphysics. Secondly, the principle of positivism states that we can not use ultimate natures (or metaphysics) as a premise for our scientific model; which is not the same as saying that there is no connection between the two of them. Ideas about ultimate natures can arise from the conclusions of the theory. And this is where the danger lies, because the scientific method itself is built from certain presuppositions; if we then try to draw those presuppositions out of our theories, we cannot be sure that we are not just engaging in circular reasoning. To conclude, based on science, that there is no ultimate nature is flawed because this is just a presumption put into some interpretations of the scientific method.

Science extracts certain mathematical observables from reality, and then draws conclusions from them. Now these qualities which can be expressed mathematically might be all there is to reality, or they might not be. The point is that science alone cannot say that its object of study is all there is, because by design it only considers those qualities. It might be able to prove that there is something beyond the physical world through the study of science, if it is found that knowledge of the material world alone is insufficient to explain and predict its nature and evolution, but one cannot prove through science that there is nothing beyond physical existence. It is like the house-builder who says that because he can make many buildings from brick, cement and wood, while rejecting everything else because he doesn't know how to work them, then brick, cement and wood are the only building materials there are. Nonsense. It is impossible to leave out metaphysics, and anyone who tries is bound to commit errors. Equally, it is foolish to study metaphysics without a good understanding of physics, because otherwise there will be nothing to ground it to the real world.

There are several reasons why metaphysics is essential: firstly, because scientific theories require certain definitions to start with which themselves cannot come from science, but thus only from something beyond it; secondly, because if we attempt to live without metaphysics, this will mean that we just uncritically accept the prevailing fashion, and if this fashion is wrong it might affect our science as well as our philosophy[1]; thirdly because it is impossible to consider ultimate questions without thinking about ultimate questions – if we restrict our thinking to only the sub-ultimate on principle, then we could be missing out on something of immense importance and intellectual satisfaction. This indeed was an error of the scholastics: they had, in terms of final causes and inherent powers, a very nice explanation of the world which worked reasonably well: but because they were satisfied with it and did not drill further in, they missed out on modern science. It would be foolish for a modern scientist to marvel at the magnificent intellectual edifice his colleagues have established, and then refuse to even think that there might be something even grander of which it is only a part.

So what are the presuppositions of classical physics? Firstly, as discussed, that its methodology is reasonable. That one can extract from experiment true qualities of the world that can be considered and extended in mathematical terms. Secondly, it presupposes a certain view of space and time which underlies, is independent of, and upon which motion is carried out. Thirdly, it presupposes that matter is ultimately constructed from individual hard substances which cannot be deformed or broken down into more fundamental particles. Fourthly, it presupposes mechanism: that a complex substance can be considered in terms of the motions of its parts and that these motions evolve deterministically according to certain differential equations. Fifthly, it presupposes that the laws of motion which provide these equations are independent of matter and of God (except perhaps as their author). Sixthly,

[1]Of course, one of the characteristics of a good scientist is to accept nothing uncritically, especially what is fashionable or the pronouncements of experts. That doesn't mean that we should not accept or reject fashion, tradition, or expert opinion: only that if it is something of importance then we should subject the strongest arguments for and against it to the rigours of logic and experiment before adopting or rejecting any view. Evidence matters more than the opinions of experts.

it presupposes that the mathematical qualities can be known definitely and simultaneously. Seventhly, that there is a real world out there, and that what we observe is not just a product of our minds and senses, but exists objectively in something external to ourselves. Eighthly, it assumes that the laws of motion are unchangeable and comprehensive, in that they cover all space and time (or at least all space and time we have observed: Newton himself would not have gone this far, having only proved that gravity applied in the solar system. However, the subsequent development of physical theory seems to require that we suppose its universality in terms of space and time).

8.2. Methodology

Newton was particularly critical of what he termed as hypotheses, defined as first principles not ultimately derived from experiment[2]. These must be resisted and rejected in every circumstance. He was not, perhaps, completely consistent in removing them from his own thinking, but he certainly came close. He placed most of scholastic thinking in this category, but also the thinking of the likes of Descartes. Newton was well aware that what he was capturing was not reality, but only a model of some of reality's features; his modern successors have not always followed his example.

Newton was adamant that any scientific thinking should both start and end in experiment. This is because what we are interested in is sensible[3] qualities. We can dream up the most beautiful mathematical or logical theories, but unless they are tied to the real world, then they are all in vain. Thus, to avoid presuppositions, the starting points of our investigation must be drawn from experiment. We measure something, and that is where we begin our reasoning. This step has perhaps been neglected; we are too prone to use theory as a starting point. In particular, it strikes many philosophers as a dangerous form of induction; as Hume showed, we cannot draw general observations from a finite number of measurements if measurement is our only guide. However, for Newton, such induction is necessary to ensure that our work is grounded in experiment. Were this part of the idea the whole process, then Hume might have a point in his criticism. But it is not; the ideas are subsequently checked and tested. For philosophers this doctrine is troubling, since philosophers deal with abstract first principles rather than derived first principles.

The middle step is mathematical demonstration, in which we attempt to draw more from our premises than we observed; to find patterns, and to reduce various diverse phenomena into a single framework. This unifying principle has indeed been one of the ongoing stories of physics. There is no *a-priori* reason to suppose that there is one single unified theory; but that has been the trajectory of physics. Mathematical demonstration is needed because only mathematics can provide the required degree of certainty. Only in mathematics are the definitions completely without ambiguity; the reasoning can be shown to be completely rigorous.

Finally, any conclusion we draw up must be tested and verified by experiment; where possible this experiment should be repeatable and repeated wholly independently numerous times before acceptance. We should not trust any conclusion of any theory, no matter how beautiful, unless it has been tested and passed that test. We should reject any theory that is invalidated by that test: possibly that theory only requires a minor modification, but if prediction and experiment differ by only an iota (being in mind their respective error bars), then that theory must be rejected until the modified version can pass its own test. If a theory is not testable, then it is not scientific. That doesn't necessarily mean that it is wrong, only that it ought not be done in the science faculty of the university. These final

[2]The next few paragraphs are a summary of Newton's views, as expressed in (Newton, 1713; Burtt, 1931).

[3]Meaning accessible to the senses. That the modern understanding of 'sensible' means *having sound judgement* shows how much Newtonian philosophy has infiltrated our thinking.

tests, of course, may also open up new principles and new phenomena, which are then fed back into the first step of the method.

It should be clear from this that Newton was not an empiricist (in the sense of Francis Bacon or Hume who made experiment our only source of knowledge) but was nonetheless thoroughly empirical. Furthermore, Newton restricted himself to only those features of the world which could be studied using this method. For everything else, there are other tools. Newton believed that in this method he had made a remarkable methodological discovery, bettering the work of Boyle and Galileo, and a vast improvement over Bacon on the one hand or Descartes on the other. Newton allied the exactness of mathematics with the relevance of empiricism; and (in the exact opposite of what occurs in most synthesises) his method retained the best features of both world-views, and diminished their weaknesses. The one weakness that he could not remove was that his method only applies to that which can be observed and considered mathematically; and thus is not guaranteed to capture every element of the physical world, let alone anything that lies beyond it.

For Newton, science was composed of laws stating the mathematical behaviour of nature; such laws to be deduced from phenomena and verifiable by testing against phenomena. It provides an exact mathematical description of the processes of the physical world. Anything else is not science.

8.3. The Classical View of Space and Time

The new physics was devoted to the study of bodies in locomotion. In the earlier Scholastic analysis, the final cause described why bodies moved, but little attention was paid to the details of how they moved. Indeed, the concept of locomotion was not well defined. With the idea of final causality abandoned, and that matter was only extension adopted, an alternative view of motion needed to be supplied, and this meant an alternative view of space and time. Locomotion had to be expressed mathematically, which meant that velocity and acceleration needed to be given precise definitions. Force, resistance, velocity and acceleration were defined in ways in which those used to think about lines, curves and solids would find appealing.

According to Aristotle, space is a property of natural bodies, it is not something which is more fundamental than them. In other words, space can only be considered if there is matter present; it is a property of matter and has no independent existence apart from matter. Similarly, time is the concept which underlies the change of matter: without change, there can be no time; without matter, there can be no change; and therefore without matter there can be no time. Empty space is thus an impossibility, and for this reason the idea that bodies are made of discrete and separated atoms rather than being continuous is inconceivable. If you have atoms, then what can lie between them? Time equally was just the measurement of the continuous transformation of potentiality into actuality; the idea that there might be temporal gaps between substances was also impossible. Correspondingly, space and time were not regarded as very important in Aristotle's physics.

The early scientists turned this around. Rather than make space and time depend on matter, they made it something which has an independent existence. If there were no matter, then there would still be the underlying space and time. Space is no more than the measure of extension; time is the measure of duration. Bodies extend across a region; they occupy a certain amount of space and last for a certain amount of time. Velocity is the measure of a how much distance is covered in a certain time. If distance and duration can be treated mathematically, then it must be possible to apply a number to them, which means that there must be an underlying measurement scale, and an underlying medium on which we can make measurements: this is space for distance, and time for duration. Since material bodies were seen as primarily in terms of extension and motion, space and time were raised from obscurity to become the key concept. The real world is no longer seen in terms of bodies changing according to their internal properties, but a world of mathematical motions

in space and time. Time is (in the eyes of a pre-twentieth century physicist; Einstein would change this) a measurable continuum; only the present exists, and the movement of time is from the past into the future, with the present that moment between the two. This means that there is nothing actual, since the present is only a moment, and everything is in a state of flux and change. Time becomes just a fourth dimension, with the only fundamental difference from space that it flowed in only one direction, from the past to the future (again, Einstein and Minkowski would take this idea and perfect it). Like space, time can be viewed in geometrical terms, this time as a line.

We therefore have to change our definition of space and time from the Aristotelian. I will state these now in full, as they are adopted in my own metaphysics. While I will try to resurrect many of Aristotle's ideas, I do not believe that his understanding of space and time can be accepted without modification.

DEFINITION 8.3.1. **Space-time**:
I define space-time as an entity that can be mapped to a four dimensional Hilbert space with the metric having a Minkowskian signature (i.e the time-time component of the metric has the opposite sign to the eigenvalues of the spatial part of the metric) and Riemann geometry upon which the laws of physics operate. Each point within this space corresponds to an event. Material objects are (naively) associated with an (open) line within space time; each electron, for example, can be described by a mapping of a mathematical function to space time; in Newtonian physics this mapping would be to a line monotonic in time (which would describe the electron's position for each point in time where the electron exists). In this way, all matter can be associated with events in space time. It is not clear whether space-time is infinite or finite. Δ

Discussion:
I have expressed this using geometrical terms, so some explanation of these terms is necessary. A Hilbert space is a way of converting a volume in an arbitrary number of dimensions into a set of coordinates. There are as many coordinates as there are dimensions, so in our case we are interested in three spatial dimensions and one time dimension, so there are four coordinates. Each coordinate is a number from minus infinity to plus infinity, and the mapping is chosen so that for each point in the volume there is a unique set of four numbers, and for each set of four numbers corresponds to a unique point in the volume. Furthermore, a Hilbert space is characterised by allowing length and angles within the volume to be easily calculated by simple relationships between three different coordinates. One example of a Hilbert space is the simple Cartesian coordinate system used to describe Euclidean (flat) geometry, where there are perpendicular axes with equally spaced ticks marking each distance along the axis. However, the concept is more general than this: one does not need to use either Cartesian coordinates or a Euclidean geometry. Why do we want to do this? The reason is simple: if we have numbers, we can apply the full power of mathematics to the problem, which is not so easy to do when dealing with a physical volume. So we set up some sort of correspondence between the volume and some numbers (some rule that converts a space-time point into a number), use mathematics to solve whatever problem we are interested in, and then convert the answer back to the point in space time using the reverse rule. This is much easier than trying to work continually with points in space-time or with pictures. $\square$

DEFINITION 8.3.2. **Energy-momentum**:
There is a Hilbert space conjugate to space-time, momentum-energy. In Hamiltonian mechanics (equivalent to Newton's mechanics), this is independent of space time, so in addition

to a position coordinate each material object has a momentum coordinate. Δ

Discussion:
Please note that this definition is wrong. The definition of energy and momentum will need updating when we consider quantum physics. It is still true that they are conjugate to time and space, but in a different sense than suggested by Hamiltonian mechanics. In classical physics, 'conjugate' means that it is defined according to Hamiltonian's principle, albeit extended into the four dimensional view of special and general relativity. Thus momentum can be derived from the dependence of the Lagrangian on the velocity, and similarly for Energy. By 'independent' of space time, I mean that the given a particular location and the laws of physics, the momentum is not constrained, but can take any value.

Energy and momentum can be described by another coordinate system – mathematically just the same as the coordinate system used to describe space and time. We can imagine a four dimensional space, where instead of time one direction represents energy and the other three directions represent momentum; then we can map these to the four numbers of a coordinate system. In some sense they are quantities invented by the physicist - what we directly observe are space and time. However, Newton realized that they are immensely powerful concepts which make calculations much easier. In particular in any impact between two bodies, both momentum and energy are conserved; so the total energy of the system remains the same before and after the collision, and the same with the total momentum. In Newtonian mechanics, the momentum is given by the mass of a particle times the velocity, and the energy by half the mass times the square of the velocity. In special relativity these formulae change, and the link between energy and time and momentum and space becomes a bit clearer. However in both of these, energy/momentum and time/space are seen as two entirely distinct spaces. Two coordinates are necessary to specify the state of a particle: the first of these is time/space; the second energy/momentum. As seen, the conservation laws are linked to symmetries – the idea that the laws of physics should be invariant under translations leads to the conservation of momentum, and so on. These conservation laws are continued into quantum physics. $\square$

8.4. God

Most of the early scientists were devoutly religious, but their view of God was different from that of the scholastics; even though there was little agreement between the early scientists themselves. The most important difference was their introduction of the "laws of physics" as an intermediary between God and nature, something having existence in its own right. For the scholastics, God acts to directly uphold the universe; the regularities in the universe are due to the constant nature of God preserving the processes of the world in accordance with the final causes of each type of being. Thus the regularities are internal to the substances and due to the direct action of God in creating and sustaining the substances. For the mechanists, physical particles are bound to obey the laws of motion, which are external to the objects, and it is the laws of motion rather than the particles themselves which are maintained by God. One of the changes when we move from final causes to a law based view of physics is how God is perceived: He goes from being the end to which all things are drawn, to merely the creator and (at best) maintainer of a universe independent of Him. Mechanism retains some form of efficient causality, so God was also referenced as a necessary first cause; however, since the mechanist version of efficient causality differs from the scholastic causality (event causality rather than substance causality), the corresponding cosmological argument also differs and is less rigorous than the scholastic version. Many viewed the laws of motion as standing on their own, with God's role to only establish them and set up the initial conditions of the universe. However, postulate an eternal universe, as many did, and through concentrating on what the laws of motion are while neglecting why

they are and why they remain constant, as the scientific method encourages us to do, and this role for God is obviously removed.

The second thing that all the mechanists were agreed on was that an understanding of teleology was not necessary for the basic understanding of motion and extended objects. Some regarded it as superfluous, if it was there, it was just contained in the laws of motion; others rejected it altogether. Of course, final causality was unanimously understood as the stoic purpose rather than the Aristotelian natural end; as far as I can tell nobody influential discussed the Aristotelian understanding of final causality in this period. With regards to purpose, the early moderns were divided. Some thought that mechanism removed any need for it, since the laws of motion explained everything; others believed it necessary to explain either biological phenomena (a God of the gaps) or the harmony of physical law (a God based on what we know). The first of these was perhaps emphasised more than the others, until Darwin showed that mechanistic principles without teleology could explain the harmony of nature. The second was neglected in later years. Thus the God of design was gradually removed from mechanistic thought. Of course, the teleologists continued to make important contributions to science: numerous early advances in biology such as the discovery of the circulation of blood, and in physics least action principles. However, the strict mechanists could easily absorb these advances into their anti-teleological world-view without recourse to the thoughts that inspired them.

Finally, particularly in Newton's writing, God had a place in correcting for irregularities in the universe. For example, Newton knew that kinetic/potential energy was not perfectly preserved: friction and inelastic collisions offered counter examples. Equally, one would expect the planets moving through the fluid of the aether to gradually slow down through friction and spiral inwards towards the sun. Thus the natural order would naturally decay, and God had the role of occasionally putting in extra energy to maintain it. Thus God was required to fill in the gaps in physics. Of course, these gaps were gradually filled in; for example it was discovered through thermodynamics that heat was another form of energy, and the internal bonds of substances could also store energy, so friction is not a violation of the conservation of energy once it is so expanded. The aether was removed from our understanding of the solar system, so the planets, if it were not for the eventual death of the sun and the odd particles of solar dust, are not prevented from continuing in their orbits forever. Plus, of course, the universe does run down, so some of Newton's examples turned out to be how the universe behaved. Thus, the irregularities that God needed to preserve disappeared, and once again the role of God diminished.

Thus although the early modern scientists were all devout Christians (or, in Newton's case, a devout Arian), and all left a clear place for God in their thinking, gradually as science developed the need for God was thought to be removed. There was, of course, still the argument from revelation, and in particular the miracles of the Bible, but it is easy to neglect this if one approaches this evidence with a mindset that has already excluded the possibility of God. Arguments against the miraculous were brought forward; theists, of course, have always find these arguments risible (certainly all those I have encountered are easily refuted[4]), but they provide the crutch that the newly godless moderns needed to dismiss historical and contemporary revelatory evidence. This is particularly appealing to those who consider the religious moral code to be an irrational straight-jacket.

The method of science also leads to a Godless position: even religious physicists make no mention of God in their scientific work. Of course, this is not surprising in any system: physics is involved in the study, explanation and expansion of regularities, and questions of why there are those regularities is separate issue. The religious man will, of course, attribute this to God; but classification of regularities can proceed without mention of God. Even the scholastic can describe the results of final causes without reference to God, even though he

[4] I discuss several of these arguments in chapter 16.

believes that God is the author and maintainer of final causality; even more can the mechanist discuss what form the laws of physics take and how particles move under their influence in isolation of the God he believes maintains them. However, because the theistic underpinnings are not discussed in the scientific method, being based on experimental observations of natural objects, it is easy for even those scientists who accept them to neglect them in their day-to-day work. Equally, any notion of God would necessarily be a non-experimental supposition of the type that Newton despised and eliminated from his method. The absence of God is not a part of the real world read out of science, but part of the framework put into the method of science, which studies what can be studied without reference to God. Whether this is everything is not a question that science alone can answer. The method of science thus does nothing to encourage theistic thinking, and, as a consequence, though this does not present an argument against God, can divert the minds of those who practice it from religious thoughts[5].

Thus it was not surprising that the role of God in modern thought was diminished, and agnosticism and atheism arose as the default position in academia in the early twentieth century, and have remained there to the present day. To my mind, if mechanism is accepted, agnosticism is the position more consistent with the basic principles of science. Atheism is, however much atheists may dispute this, a positive statement, and one not based on any empirical evidence. Atheists may present their view as a denial of the numerous world-views that include deities of some form or another, but this is the same as a statement that one of the possible world views that do not contain deities is correct. This is, of course, a positive statement. But what evidence can be presented for this position? Some have, of course, tried to find a contradiction in the very definition of 'God', but when discussing 'God' in general, this seems to be foolish, since there is enough flexibility in the term to create a loophole against any such argument. Most of these arguments I have seen have taken Kant's definition as a starting point for their attacks[6], but even if successful this does no more than show that Kant's God is self-contradictory, which still leaves Newton's God, Boyle's God, Augustine's God, Mohamed's god, and the numerous other gods which have been or could be imagined. That, of course, assumes the arguments are successful; I do not even believe that they stand against the God of Christianity, let alone all of the possible alternatives.

The other possibility is to say that the concept of God invalidates our experimental evidence; but that cannot be, since experimental science is not comprehensive, and does not cover all possible knowledge, and God certainly cannot be cornered any number of limited controlled experiments. Finding arguments to show that something does not exist is immensely harder than proving it does exist. The lack of any God in science (if there is such a lack) is more a statement about the limitations of science rather than the limitations of God. Certainly God is consistent with the mechanical universe, since so many of the early mechanists believed in God.

The key step in the removal of God from the mechanistic world view was (I believe) first stated by Laplace (see (Plantinga, 2011), chapter 3). For example, if we consider classical mechanics to be reduced to the conservation of energy and momentum, then properly stated Newton's expression of physics doesn't say that energy and momentum are conserved, but that energy and momentum are conserved in a closed system. If we consider a box in which we perform our study, then energy and momentum are only conserved within the box if that box is entirely cut off from the outside world: if not we can safely transfer energy or momentum into the box even within the scope of Newton's laws. Newton's second law states that acceleration is force divided by mass; but only if we have accounted for all the forces acting on the body in terms of (say) gravitational and electromagnetic interactions with other matter. The question then is whether the universe as a whole is a closed system;

[5]This is, of course, true for just about every area of work except perhaps politics, philosophy and the clergy.

[6]I will discuss those arguments which I consider most worthy later.

whether external forces only come from interactions with other material objects. This additional assumption is not scientific; it cannot, for example, be proved by experiment since any experiment we conduct might in principle relate to a point where there is no outside intervention; even if it holds the vast majority of the time we cannot legitimately extrapolate to say that it holds all the time, particularly if we leave out those examples where it is claimed to have broken down from our sample. The idea that the universe is closed is therefore a metaphysical assumption – and in important one for the question of God, because if it is assumed that there is nothing outside the universe which can enter into it, then direct divine influence is forbidden. People were led by this assumption to claim that God could not intervene in the universe, and to claim this as a conclusion of scientific law; however it is not a conclusion of scientific law but only of this additional assumption added to classical mechanics that the universe is closed.

But what of the teapot in orbit around a distant star? Surely I would reject this, despite the lack of evidence saying that it isn't there? Or the Poltergeist inhabiting some Irish castle? So if I reject these ideas because there is a lack of evidence for them, surely one can reject God on account of a supposed lack of evidence for God? But examples of this type are flawed. Firstly, and most importantly, they commit a category error. God, as classically conceived, is nothing like a teapot or a poltergeist. The classical definition of God is a being of pure actuality; whereas teapots and poltergeists, if they exist, are a mixture of actuality and potentiality, so they by themselves lack an explanation of why they are in the state they are in; whereas with God, no explanation is required, since He cannot be in anything other than what He is. There is no resemblance whatsoever; so the analogy cannot stand. Secondly, people have claimed to have evidence for God, while I know of no evidence for a teapot around a distant star. Of course, atheists would reject those claims for evidence, but then, atheists are just as capable of making an error in the rejection as an (equally intelligent) theist is in making the claim for the evidence; more so, since atheists are usually less knowledgeable about the particular branch of theism in question than that theist is: the theist has made it his life's work to study and understand it, while most atheists have not[7]. True, the theist comes in with numerous presuppositions, but so does the atheist, and again, the contemporary theist has had to consciously reject the presuppositions of the modern world, while the atheist may well have given the matter less thought since he agrees with the presuppositions inherited from the prevailing intellectual culture. Thirdly, I do not claim to reject such things; I am very much a teapot agnostic, as I know little about the evidence for or against it's existence, nor even on why that teapot should make such a big difference to my life.

Nor does the objection that there are many different possible Gods out there invalidate that one of them could be correct. After all, there are many different forms of atheism, but no atheist would use this to claim that all of them must therefore be wrong. The different Gods, or at least those most plausible, differ mainly in the finer details; the differences may just be a matter of imprecision rather than inaccuracy, and no scientist would regard imprecision of knowledge as a barrier to believing.

The religious man claims to accept God, partly on philosophical grounds, and partly on the grounds of observed evidence (whatever revelation he accepts). The atheist rejects God only on the basis of philosophy, since there is no observation against God, and he rejects all claimed evidence for God. But this is just a denial of the scientific method; which, while requiring both a theoretical construct and observation, first and foremost states that no

[7]Obviously, I do not wish to over-generalise here. Firstly, I am solely speaking of well-informed theists and well-informed atheists, that is those who have made a detailed study of their own beliefs including answering objections against them, which removes the vast majority of people in both camps. Secondly, there are obviously numerous atheists who have studied theism in detail, and understand it at least as well as the best theists. My argument is that the well-informed theist is necessarily an expert on his own particular brand of theism, including the main arguments against it; the well-informed atheist need not be an expert on all forms of theism.

philosophical argument should stand unless backed up by observation, nor can you override observation with metaphysics or authority. The agnostic, on the other hand, takes a more honest position; stating that the lack of convincing evidence either way implies that he should not form an opinion either way. Of course, practically the agnostic lives the same life as the atheist.

The religious, of course, responded. There were issues in the philosophy of mind and epistemology, so these were cracks that God could sneak in again. Thus some pleaded for the existence of a divine sense within us; some used the ambiguities of language and sensation to cast doubt on the comprehensiveness of science with regards to the physical world; some postulated a division between the real world (which contains God) and the world of our ideas (which contains our representation of physics); or some said that Church is is just what respectable people go to, isn't it? therefore God; and others postulated that there *really must* be something out there to explain it all, mustn't there? therefore God. Atheists found such responses (which can obviously be expressed more strongly than as phrased here) risible; frankly, so do I. They add no real explanatory power and are not based on experiment nor even (if accepted clearly miraculous) religious tradition or observation, so cannot even reasonably claim to have come from God. While it would take more words than I can spare here to refute them, I do not take such possibilities seriously.

However, the removal of God from the mechanical universe has led to an undermining of that same universe, because the founding presuppositions of the mechanical world-view were built on either theism or deism. What are the laws of physics? Why are they universal, constant throughout all time and space? Why are they differential equations, built around forces and acceleration? If they are external to matter, then what connects them to matter that makes matter obey them? There is also the mind-body problem, raised by Descartes as a consequence of his philosophy, and still not satisfactorily answered within the context of mechanism[8]; and the is/ought problem, raised by Hume, and whose only solution consistent with the mechanistic presuppositions is the unpalatable Nietzschean denial of morality. Originally, all these ideas were first postulated and defended on account of the relationship between God, the laws of physics, and the material universe. Yet God has been kicked away, and nothing put in His place. This leaves the mechanical world without foundations, or any reason to accept it baring the success (in its limited field) of the science built upon it. But this success is no argument for the mechanical world view, since it could well be that other world-views could also give similar, and maybe even better, results. Such success is a necessary condition (or to be more precise, the possibility of such success is necessary – one can imagine a world-view which could in principle lead to a successful scientific theory, but nobody happens to have done it yet, either because it is too difficult or because nobody has really tried; such a world-view cannot be ruled out on these grounds), but not sufficient.

So God does not have much of a place in the mechanistic world-view. The laws of Newtonian physics are seen to be fundamental in themselves; they operate independently of God. At best, God is seen as just the architect of those laws; who set the universe in motion and can do nothing else. Possibly one may hold that God occasionally intervenes, breaks the laws of physics, but there are arguments against this. The universe is capable of acting and evolving on its own without any divine intervention. This position is known as deism.

[8]The mind-body problem is actually several problems. In the context of Cartesian Dualism, it relates to the relationship between the physical (and mechanical) brain and the immaterial (and non-mechanical) mind. If a mechanistic picture of the mind is accepted ahead of Descartes' dual picture, then it is how can brute matter, with no inherent directionality, universality or meaning give rise to thought which we observe has meaning, a direction, and can ponder other objects and study their relations? Then there is the difficulty of free will, which a purely mechanical model of the mind cannot account for, and yet is something we directly experience, or at least are fated to believe that we directly experience. Scholastic philosophy does not have these problems, because it claims that matter does have an inherent directionality (the final cause) and there is some part of objects which can be grasped intellectually (the form): it is precisely the denial of formal and final causality which led to the mind-body problem; yet this denial is the one thing that connects all forms of the mechanical philosophy (Jaworski, 2016).

It is thoroughly incompatible with Christianity and the other theistic religions. The role of God is reduced simply to explain why things are as they are, and as scientific knowledge increases, so does the role of God diminish; until we are left with nothing at all. There is one important note to observe here: most atheistic arguments I have seen implicitly presuppose mechanism, or some similar enlightenment framework; and within this framework, these arguments appear to be quite reasonable. They do not, however, have the force which those wielding them believe, because they implicitly or explicitly attack deism, rather than theism. The Christian also disagrees with deism, but for a very different reason: we reject as unscientific the underlying mechanistic philosophy.

Is it possible to reconcile enlightenment philosophy (which has mechanism as one of its streams) with monotheism? I do not believe so (although I cannot prove this), for two reasons. Firstly, it has been tried and failed. Liberal (or 'progressive') theology (which I define as an attempt to ground a theological system on enlightenment rather than classical philosophy) started off nobly, to try to make Christianity accessible to the modern reader, but as it has developed, it is now obvious as an intellectual dead end: neither rational, nor recognisable as Christian. Gradually every distinguishing doctrine of Christianity has been stripped away, until (in its least unreasonable form) one is just left with 'atheism in ceremonial robes.' Jesus cannot be seen as the incarnation of God (God made flesh) - this requires the virgin birth, which is seen as scientific nonsense and one of the first doctrines to be denied, so he is just seen as a man who was more aware of the inherent 'God sense' within all of us than most. God cannot be understood at all by human intellect; we cannot make definite statements about Him/Her/It and all historical experiences of the God-sense (whether Christian or not) are just mystical experiences useful for that time but not necessarily ours; and as the Church grows its awareness of the ineffable divine increases, and so it may discard primitive views for the latest development discernible through experience, supposed reason, and the latest cultural (even atheist) fashion. Such a viewpoint became dominant in many European and North American Churches in the twentieth century; firstly adopting modernism with its denial of the miraculous (and thus much of the core of Christianity) and then post-modernism, with its denial of a knowable objective reality and raising subjective self-identity as though it were the only sure principle. Thus liberal theology is left with just empty (though well meaning) platitudes. It is not a religion that will survive for much longer, as it has no power and nothing to offer; but it has distracted and devastated much of the Christian witness in the West, consuming the energy of and destroying every church it infects. The flaw, I believe, comes from the start; the premise that it is possible to merge enlightenment philosophy with Christian theology. Its history provides a *reducto ad absurdum* demonstration that that premise is flawed.

My second objection is more theoretical: enlightenment philosophy is defined by the presumption that there is no final cause; i.e. no individual material objects has inherent tendencies. However the theology of all the major theistic religions believe, if nothing else, that we were created by God for a purpose (which leads to natural tendencies within His creation); that God is specifically interested in us. That each of us has our own calling, and goal, in Christianity to be conformed to the character of Christ and to live in fellowship and to the glory of God, and, indeed, everything in creation is designed to fit in with these various purposes (note I am not saying that a particular asteroid in a distant galaxy is there for our benefit; merely that the principles that allow that asteroid to exist, and make it likely that such asteroids do exist, are the same principles which allow those objects which we do depend on to function). That, because of our corruption and imperfection, this purpose is impossible for us to attain by our own strength and powers, but through the instruments offered by God, most importantly the sacrifice of Christ and union with Him, and no merit of our own, God can remake us to fulfil what we ought to be and need to be to find true happiness, if we accept and trust Him. This purpose from God would reflect itself at a microscopic level as the display of tendencies in matter towards certain ends, those ends

adding up to some macroscopic effect which has the potential of being consistent with the fulfilment of God's purpose. In short, if a God exists who in any way resembles the God of the major theistic religions, then there is final causality in the world. Yet liberal theology both demands the existence of God and builds on philosophies which deny final causality. Thus, there at least appears to be, a discrepancy between enlightenment philosophy and the theistic religions at a foundational level.

Equally, enlightenment philosophy, and thus also liberal theology, denies the possibility of form. Yet God has knowledge about us – not the modelled, approximate or vague awareness we possess, but a true, complete and accurate knowledge – which means that some aspect of our nature must be an abstract quality accessible to the intellect. We fall short of God's standards for us, and God desires us to improve in character and virtue while remaining the same being, knowing not only what we are but what we ought to be, which means that that abstract description of us must contain different states, some actual and others potential. We have reached the scholastic concept of form. Thus to deny form and accept God entails a contradiction, one which no liberal theology can avoid. Thus liberal theology is no more than an irrelevant distraction: the debate is between enlightenment atheism and theism built on classical philosophy.

8.5. Causality

The classical physics view of causality is obviously very different from the scholastic/Aristotelian view. Firstly, let us consider the four causes.

That the world is seen as deterministic is seen as removing the need for final causality; ultimately it is just efficient causality expressed in a different way. The removal of final causality is one of the cornerstones of modern thought.

Formal causality is also severely curtailed. The arrangement of matter certainly plays a role in classical physics, since it is by different arrangements and orientations of the fundamental corpuscles that more complex forms of matter emerge; and this is part of the scholastic notion of form. However, form goes beyond that, and states that new properties can emerge when we combine different elements. A water molecule, according to classical physics, is just two hydrogen atoms and an oxygen atom, together with the electric fields binding them together, since a body is just the sum of its parts. However, the scholastic concept of form allows for something new to emerge; a water molecule is something different from the sum of its parts. Equally, the idea that something can be in different possible states and yet be the same thing is absent from classical physics, and consequently also from all the different forms of enlightenment philosophy (in my view, one of the biggest weaknesses of modern philosophy). There is no concept of potentia in classical physics. Thus the scholastic vision of form is severely curtailed in classical physics, and generally abandoned by the philosophers.

With regards to material cause, some of the philosophers abandoned the concept of matter entirely, believing that the fundamental building blocks of the universe are instead some geometrical object. Most, however, would not go this far. Objects are still made of matter, but the type of matter is not seen as so important, since all matter obeys the same basic laws. So while we can say, 'This wire is made of copper' or 'This knob was constructed from brass,' the fundamental level of physics goes beyond the copper and brass to the corpuscles themselves, and at this level there is no real distinction between one form of matter and another. Thus the role of material causality is also diminished in classical physics.

Then we come to efficient causality, and it is commonly held that this alone is emphasised in classical physics. However, it is important to remember that the Aristotelian efficient causality refers back to a substance (or, more precisely, a substance in a specific state), and I will continue to use the term efficient causality to describe this. However, in the mechanical world view, motion is transferred through forces and collisions. Thus

the cause of a being's motion is often expressed as the event that created its motion; and events also have causes in terms of the various particles which collided to create that event. The corpuscles themselves, of course, are indestructible so cannot be created and destroyed (except by God in a moment of creation), and thus cannot themselves be caused. I describe this concept as event causality.

This new view of causality led to various issues in philosophy. One thing that has been noticed is that classical physics as such rarely talks about causes. Instead, we write down differential equations and solve them. However, there is still an idea of ordering in time, either meaning (in Newtonian physics) only those things in the past can influence the present, or (in special relativity) no information can pass outside the light cone; so the only objects we can be aware of are those connected to us by light beams or slower sources of information, and the only objects which could be aware of us are those which could receive light signals from us. This is important as one of the consistency checks of quantum field theory is that this property is restored. But otherwise, it is certainly true that a classical physicist will hardly ever, if at all, use the concept of causality in his calculations. Does this mean that causality is an illusion? Not entirely. Modern physics can only study what can be mapped to an algebraic space, and causality is not a mathematical concept. We should not therefore expect to see the concept in the equations, but that doesn't mean that it is not present. Causality, after all, arises from the belief that the universe is rational. The belief that at least some aspects of the future state of the universe can be explained (which aspects and the scope of the explanation depending on the type of causality being considered) in terms of the past state of the universe. This idea is very much a part of classical physics. Classical physics assumes causality; its equations are a description of how a particular type of causality is realised.

I should, however, emphasise that we ought not demand that we leave Aristotle's metaphysics unchanged; I would rather take him as a starting point than an ending point, and if convenience or physics demand that certain aspects of his thought need to be modified, then so be it. What I want to ask is whether modern physics demands a complete rejection of Aristotle, or whether a minimally modified version of Aristotle's ideas can be fitted match to classical physics to see if he can provide an alternative to the dominant mechanistic view.

The first principle of (efficient) causality is that nothing is moved except by something external to it. This seems to conflict with the principle of inertia, which states that things continue in motion unless acted on by an external force; however the conflict is only because there are conflicting definitions of the word *motion*. In the principle of causality, motion means movement from one state to another, or one potentia to another. A state means something that is stable in time; in Aristotelian physics this is a body at rest in its natural position. In Aristotelian physics, a state means a having a definite location, regardless of the velocity. In Newtonian physics, we can define a state as having constant velocity, regardless of the position. This continues into general relativity, with the 'states' now described by the geodesics in four dimensional curved space time, with the benefit that the geodesic incorporates all information about the external gravitational 'forces'. In all of these there is some equivalent of Newton's first law, in that one can only change one's state if acted upon by an external force. But this is nothing more than the principle of causality adopted into this new understanding of motion, since we are rewriting the principle of causality as 'No inertial state can change (i.e. being moved in the original language) except by something external to it.' Thus the principle of causality remains in classical physics, encoded in Newton's first law. This means that the concept of (efficient) causality is still there, only different words are used to describe it. We still, of course, have the problem of action at a distance for classical electromagnetism; the early mechanists believed that motion could only be transferred by collisions, and this is easy to incorporate into such a view of causality. The mutual forces action at a distance of classical electromagnetism seem to be harder to

account for; however since electromagnetism is due for a considerable overhaul when we consider quantum effects, I will leave this matter here without further discussion.

A further important clue to causality in classical physics is the conservation of energy and momentum. Like the first law of motion, these are key concepts that survive the transition from Newtonian physics through to general relativity. What this tells us is that energy and momentum are never created and destroyed. This means that if you encounter an object with a certain amount of four-momentum, it could not have emerged from nothing, or from something with less four-momentum than it does. In other words, every 'lump' of energy had to arise from something else with more energy breaking down into that lump and something else to carry the rest of the energy, or from two or more corpuscles colliding and fusing together into one or more resultant objects. For every lump of energy, there has to be another lump of energy proceeding it. What cannot happen is that it emerges from nothing. Furthermore, effects can only be created by causes with precisely the right amount of four-momentum. This is the second scholastic principle of causality, that the effects are proportionate to the cause.

In other words, if we think of causality in terms of packets of energy and momentum, rather than substances in or out of motion, the basic ideas of causality continue to be fulfilled in classical physics. The only difference is that the two principles of causality are rephrased as Newton's first law, and the conservation of four-momentum (which is drawn from Newton's third law). This is, of course, good: Newton's law and the conservation equation are more precisely and mathematically defined. The difference between them and the older causal principles is that the causal principles are fuzzier: there are numerous different ways in which these causal principles might be expressed in different formulations of physical laws, of which Newton's formulation is just one. This means that if the causal principles are true, then it doesn't follow that Newton's laws are correct, since there are other ways of satisfying them. However, if Newton's laws are correct, then the principles of causality are also true.

This idea of causality applying to packets of energy and momentum is further complicated when we turn to special relativity, and the famous equation $E = mc^2$. This equation tells us that mass is a packet of energy; and since every physical body in classical physics is a mass, then the principles of causality must also apply to physical objects. The difference, of course, is that energy need not only be contained in masses, but also other effects such as waves and fields: these are also packets of energy, albeit extended rather than localised. Thus we need to extend the principle of causality: no longer is it just linking one substance (or material object) to another; it now links material objects, waves and electromagnetic fields to other material objects, waves and electromagnetic fields. Waves and particles are straight-forward to include into the concept of efficient causality, because they are localised packets of energy, analogous to the localised Aristotelian substances; fields are a little more troubling because they are truly extended and would thus require a substantial reworking of the Aristotelian idea of efficient causality, but again it is not worth pondering this too deeply until we have discussed the quantum theory of electromagnetism. However, leaving aside this caveat, one can see how the scholastic concept is still present in physics, only made more precise and detailed.

We can also take another look at final causality; let us concentrate on Newtonian gravity. Now we cannot say that the final cause of a force is to accelerate an object, since final causality only applies to substances and force is not a substance. However, we can summarise one of the final causes of any piece of matter through the two statements: 1) Every piece of matter generates a gravitational potential that is proportional to its mass and inversely proportional to the distance from that matter; and 2) Every piece of matter in a gravitational potential experiences an acceleration equal to the differential of that potential. We can put these two statements together in two different ways, firstly either by treating the gravitational potential as a substance in its own right (with the different

possible gravitational fields representing its different potentia), in which case the first of these statements is a direct statement of final causality, and the second can be re-written as the gravitational potential has the final cause of initiating an acceleration in any matter; or secondly, if we are not willing to treat the gravitational potential as a substance, then by just eliminating the middle step and saying that each body of matter has the final cause of accelerating towards every other piece of matter, with the magnitude and direction of the acceleration given by Newton's law. Since acceleration, in this notation, is just movement from one potentia to another, and final causality is a description of which potentia can be actualised in given circumstances by a substance, then clearly it is reasonable to discuss this tendency of matter in terms of final causality. These two approaches are equivalent; I feel the second is a little truer to the spirit of Aristotle, but the first is a little less cumbersome and more easily generalised and I have no objections to modifying Aristotle's metaphysics when convenience or physics demands it. This approach can also be easily generalised to general relativity by stating that every mass has the inherent power of curving space time, and that matter has the final cause of travelling along geodesics of curved space time (i.e. the final cause of matter at one point along the geodesic is a similar object at the next point along the geodesic); such motion represents one potentia of the substance. To move from one potentia to another would require some non-gravitational interaction.

So instead of saying that masses have the effect of inducing a gravitational force, we can say that masses have the power of curving space time. These two statements have exactly the same meaning, only expressed in different words. So why discuss final causes and powers when we can discuss things in the mechanistic language of forces and inertia? Well, why discuss things in terms of forces and inertia when we can use the language of final causality and potentia? But isn't saying that masses have a final cause of generating a gravitational potentia a hindrance to discussion? No more so than Newton's statement that masses induce a gravitational force. We are still left with the question of why they do so, only in Newton's language it is perhaps even less apparent than in the discussion of final causality. Of course, as we refine the theory of gravity, moving beyond Einstein's theory to some quantum theory of gravity we may find the ultimate answer of why masses have this particular power. As we do so, we will no doubt have to modify our statements concerning the final causes to match the new layers of physics. But to say that we have misidentified the final causes is not to say that we have misidentified that there are final causes.

Thus the contention of Bertrand Russell and others that physics has eliminated even efficient causality because physicists no longer talk about causes is false: physicists always have discussed efficient causality (baring in the short period between the abandonment of scholasticism and the acceptance of Newton's laws) and final causality, but have confused the philosophers by calling them by different names.

However, this confusion has had the serious effect on philosophy that most people, when they think of causality, only think of event causality rather than efficient causality. This is inspired by the concept that motion is caused by collisions, i.e. events in space time where motion changed. The principle of event causality is that every change of state is caused by some event; and every event is caused by the conjunction of various states. This dual principle is necessary to maintain the chain of causes and effects; if either one of the links is broken, then we are left without any explanation under event causality. When using event causality, we have to define a state as a definite momentum and a definite position, since both of these are required to predict where and when a collision and thus an event will occur. This means that event causality is not the same thing as efficient causality, which focuses only on energy and momentum, and if we include position in the definition of the state then the principle of causality is no longer equivalent to Newton's first law. Classical physics can thus be thought of either in terms of event causality or efficient causality; however the distinction between the two has not always been made. Event causality is often seen as being more useful in performing calculations, and perhaps for this reason most modern

philosophical discussion has concentrated on event causality rather than efficient causality. Event causality has no place in the scholastic efficient causality. However, it can be treated as part of one particular expression of final causality, with the only difference being the introduction of an intermediary step, the event.

In scholasticism, we would state that the final cause of the two particles coming together in a collision is the resultant effects. This fits in with the idea that the state of the particle being just described by the momentum, since scholastic final causality allows for many different possible effects arising from the same state. The final cause is conditional on various circumstances being satisfied. In this case, those differences are related to the unspecified position of the particle. The efficient causes of a substance are fixed by the past history of the substance. The final causes are conditional on external circumstances; it is possible for the locomotion of a ball to be diverted in a particular way, but only if it is struck by another ball with exactly the right momentum and trajectory; if the details of the collision were different, then the change in the locomotion of the ball would be different. Thus final causality, even in classical physics, is conditional on external circumstances rather than determined, while efficient causality is determined. Therefore they are not simply two expressions of the same thing, and neither of them are redundant.

So both final and efficient causality are indeed present and distinct in classical physics, if we define states as only describing definite and momentum. However, this is troubling, because in classical physics location is independent of momentum and is part of the description of an object. So the concepts fit, but not as snugly as the scholastics would like[9]. Once again, however, I will not work to resolve this problem, since the issue changes once we move to quantum physics. However, because of the change in language and emphasis, and the issues I have left unresolved, most modern philosophers have not noticed this, and just talk about event causality as though it is the only relevant type of causality in physics, if they think causality is useful at all; and this talk has filtered through into wider academia, including the science departments.

8.5.1. Humean Causality

I should mention one additional discussion of causality in modern philosophy, that of David Hume. Since many modern criticisms of theism date back to Hume, and Hume's views on causality are central to his ideas, I shall discuss this in some detail.

Hume starts by reflecting on the nature of relation between different objects. He emphasises first of all his empiricism, that concerning material objects we can only be certain of that which we perceive with our senses. The connection between objects is known as causation.

> According to this way of thinking, we ought not to receive as reasoning any of the observations we may make concerning identity, and the relations of time and place; since in none of them the mind can go beyond what is immediately present to the senses, either to discover the real existence or the relations of objects. It is only causation, which produces such a connexion, as to give us assurance from the existence or action of one object, that was followed or preceded by any other existence or action; nor can the other two relations be ever made use of in reasoning, except so far as they either affect or are affected by it. There is nothing in any objects to persuade us, that they are either always remote or always contiguous; and when from experience and observation we discover, that their relation in this particular is invariable, we always conclude there is some secret cause, which separates or unites them[10].

[9]See (Feser, 2012) for a discussion of how Aristotelian causality fits with the principle of inertia; and also my brief discussion below in section 8.7.

[10]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section II.

Note here Hume's implicit denial of final causality or the Aristotelian notion that objects have implicit powers or abilities: there is nothing in the objects of themselves that persuades us that they are remote (i.e. not connected to each other via causation) or contiguous (i.e. next to each other); at least, as we cannot directly observe such a property, we cannot be sure that it is there. Hume next goes on to consider whether we can have any confidence that the same object can continue to be itself over a course of time, and again he casts empiricist doubts, but affirms that we have reasons to accept it:

> The same reasoning extends to identity. We readily suppose an object may continue individually the same, though several times absent from and present to the senses; and ascribe to it an identity, notwithstanding the interruption of the perception, whenever we conclude, that if we had kept our eye or hand constantly upon it, it would have conveyed an invariable and uninterrupted perception. But this conclusion beyond the impressions of our senses can be founded only on the connexion of cause and effect; nor can we otherwise have any security, that the object is not changed upon us, however much the new object may resemble that which was formerly present to the senses. Whenever we discover such a perfect resemblance, we consider, whether it be common in that species of objects; whether possibly or probably any cause could operate in producing the change and resemblance; and according as we determine concerning these causes and effects, we form our judgement concerning the identity of the object.
>
> Here then it appears, that of those three relations, which depend not upon the mere ideas, the only one, that can be traced beyond our senses, and informs us of existences and objects, which we do not see or feel, is causation. This relation, therefore, we shall endeavour to explain fully before we leave the subject of the understanding[11].

So now Hume begins his discussion of causality; once again he considers primarily the impression on our senses as the primary clue in the investigation. He denies that it can be related to the properties of the objects themselves, as whatever property he finds, he can find another object which lacks that property but still comes under the rule of cause and effect. By properties, of course, he means those properties we can directly or indirectly sense, such as colour, size, shape, mass and so on. The Aristotelian conception, of course, is that all (contingent) objects are subject to efficient and final causes as a consequence of their coming into and departing from actual and potential existence. It is that they are an existent union of matter and form that enables them to have this relation – and this is shared by all material objects. The particular possible efficient and final causes subsist within the nature of the object. Hume's first objection to causality is thus a straw man; not something any proponent of causality would claim. However, his conclusion from this observation, that efficient causality is primarily a relation between different objects, is not something I would dispute.

He starts by observing that causes and effects tend to be contiguous rather than distant (through in classical physics, gravity and electromagnetism offer obvious counterexamples; admittedly ones overruled in quantum field theory). Secondly, he affirms priority in time: the cause always proceeds the effect; indeed the effect, at first, is perfectly co-temporal with the cause. This, for the Aristotelian is only partial; it is true only for efficient causality (and in particular final causality is directed towards future events), but as Hume's discussion is limited to efficient causality, this should not be seen as a drawback. He also notes that there is not a sole cause (no spontaneous decay), but only when the cause interacts with an external principle which allows the actualisation of the potential which it secretly possessed.

But this is as far as Hume can go:

[11]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section II.

> Having thus discovered or supposed the two relations of contiguity and succession to be essential to causes and effects, I find I am stopped short, and can proceed no farther in considering any single instance of cause and effect. Motion in one body is regarded upon impulse as the cause of motion in another. When we consider these objects with the utmost attention, we find only that the one body approaches the other; and that the motion of it precedes that of the other, but without any sensible interval. It is vanity to rack ourselves with farther thought and reflection upon this subject. We can go no farther in considering this particular instance[12].

This is as far as considering cause and effect as relation takes him. Hume finds this state of affairs most unsatisfactory; there must, surely, be a *necessary* connection between cause and effect. Stuck for the moment, he starts turning his attention elsewhere, until he has examined various neighbouring fields and looked at various questions.

> First, for what reason we pronounce it necessary, that every thing whose existence has a beginning, should also have a cause?
>
> Secondly, Why we conclude, that such particular causes must necessarily have such particular effects; and what is the nature of that inference we draw from the one to the other, and of the belief we repose in it?
>
> I shall only observe before I proceed any farther, that though the ideas of cause and effect be derived from the impressions of reflection as well as from those of sensation, yet for brevity's sake, I commonly mention only the latter as the origin of these ideas; though I desire that whatever I say of them may also extend to the former. Passions are connected with their objects and with one another; no less than external bodies are connected together. The same relation, then, of cause and effect, which belongs to one, must be common to all of them[13].

With regards to the first question, he notes that most have taken it to be self-evident, giving the matter little attention. But Hume disputes that it is so certain.

> All certainty arises from the comparison of ideas, and from the discovery of such relations as are unalterable, so long as the ideas continue the same. These relations are resemblance, proportions in quantity and number, degrees of any quality, and contrariety; none of which are implied in this proposition, Whatever has a beginning has also a cause of existence. That proposition therefore is not intuitively certain. At least any one, who would assert it to be intuitively certain, must deny these to be the only infallible relations, and must find some other relation of that kind to be implied in it; which it will then be time enough to examine.
>
> But here is an argument, which proves at once, that the foregoing proposition is neither intuitively nor demonstrably certain. We can never demonstrate the necessity of a cause to every new existence, or new modification of existence, without showing at the same time the impossibility there is, that any thing can ever begin to exist without some productive principle; and where the latter proposition cannot be proved, we must despair of ever being able to prove the former. Now that the latter proposition is utterly incapable of a demonstrative proof, we may satisfy ourselves by considering, that as all distinct ideas are separable from each other, and as the ideas of cause and effect are evidently distinct, it will be easy for us to conceive any object to be non-existent this moment, and

[12]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section II.

[13]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section II.

> existent the next, without conjoining to it the distinct idea of a cause or productive principle. The separation, therefore, of the idea of a cause from that of a beginning of existence, is plainly possible for the imagination; and consequently the actual separation of these objects is so far possible, that it implies no contradiction nor absurdity; and is therefore incapable of being refuted by any reasoning from mere ideas; without which it is impossible to demonstrate the necessity of a cause.
>
> Accordingly we shall find upon examination, that every demonstration, which has been produced for the necessity of a cause, is fallacious and sophistical[14].

Hume goes on to dismiss various reasons for supporting this principle; but his arguments need not detain us for the arguments he attacks are not those which I would use. Let us first of all examine the argument above. In the second paragraph of the quotation, Hume commits a clear fallacy: he confuses the intellect and imagination. Our imagination may well be irrational; it is certainly possible for men to *doublethink*, to hold mutually contradictory ideas at the same time, as has been proved by numerous people. The laws governing our imagination differ from those, accessible to our intellect, which govern the real world. That we can imagine, for example, a rabbit appearing out of nowhere does not demonstrate that rabbits can in fact appear out of nowhere. I can imagine a creature which is the same as a pig except that it has the wings of a sparrow growing out of its back; and that it only needs to flap those wings to be able to soar into the clouds. I can imagine it; but reason tells me that the wings combined with a pig's muscles could never generate enough up-thrust to overcome the force of gravity.

There are several other mistakes in his argumentation; for example, he defines distinct objects as those which are separable, and then states that since cause and effect are distinct they must be separable. He makes the mistake of using the word *distinct* in two different senses. Sometimes, to be distinct means separability – a cow and and a chicken, for example – but otherwise things can be distinguished but not separated, or separable in some respects but not others – such as the two ends of a piece of string. Even if we cut the string, each end would still have another distinct end co-joined to it. In his first premise, Hume uses the word *distinct* in the first sense, but cause and effect are only distinct in the second sense: every cause by definition implies an effect, and every effect a cause. One may cut the chain of instrumental causes, but the chain will always still have two ends linked together.

For an object to be separable from something else, their joining must be accidental to the nature of the being, so in some potentia (or no potentia) they are co-joined together, and in others they are separated. For it to not be separable, their relationship must be essential to a beings nature, i.e. it applies to every potentia of the being. Some properties of a being, such as its location, are accidental, while others are essential to its nature. So by saying that something is separable from a being, we mean that it is an accidental property; to say that it is inseparable is to say that it is an essential property. The question is whether a particular being's efficient cause is accidental or essential to its nature; and it is clear that it is an essential feature, since every potentia emerged from the same cause, albeit through other potentia as intermediate steps. Thus while an effect may be separated from its cause in many respects, that the cause was the cause of the effect cannot be changed.

Consider a strawberry (an effect) and the plant it came from (a cause). Hume claims that cause and effect can be separated, and at first glance this seems reasonable, because we can pluck the strawberry from its plant, and study it in isolation. But we cannot study everything concerning the strawberry in isolation, that is everything essential to its nature. We would show that the strawberry is separable from its plant in some effects, but cannot show it in every respect, because by isolating it we limit our studies to only those essential properties of the strawberry which can be considered in isolation. We cannot

[14]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section III.

say that there are additional essential properties of the strawberry in this study, but nor can we rule them out: the question one that by selecting this particular methodology we leave ourselves incapable of answering. But other, non-empirical, methods can provide the answer. Every strawberry emerges from some particular plant – which we derive not from common observation but from knowledge of the strawberry's form and by unpacking its DNA. If we consider strawberries in general, then which plant was its efficient cause is accidental to the species, since some strawberries would come from one plant and others from another; but once we come down to an individual strawberry the efficient cause is essential to the nature of that fruit, since it is shared by all the potentia of the strawberry. Therefore the cause is not separable from the strawberry's nature. To consider something fully requires acknowledging all its essential properties and features. We can, of course, look at the strawberry and not think about the plant it came from, but that is because we only understand it partially. But the nature of the strawberry itself is not constrained by our lack of knowledge concerning it. Our picture of the strawberry does not take away that it is of its essence, one of the things that distinguishes it from other strawberries, that it came from one particular plant. Hume's mistake, then, was to think that because he can see that the strawberry is separable from the plant in some respects, that it is therefore separable from its cause in every respect. Whether that is true is precisely the point under contention. Hume's argument, therefore, is a simple case of begging the question, and hiding it (perhaps even from himself) by defining a word differently in each of his two premises. This basic mistake in his reasoning completely invalidates Hume's conclusion.

The principal reason that physicists would hold to the principle of causality is due to conservation laws. Here we have certain features – energy, momentum, electric charge, and mass in Newtonian physics which must remain constant in time; indeed we also have conserved currents, where the flow into a space-time volume must equal the flow out of the volume. These follow from symmetry laws applied to a Lagrangian method of physics; to deny such conservation laws, one must either deny that the symmetries apply (which in some cases seems to me to deny that there are physical laws, so fundamental are the symmetries), or that physics can be described by a Lagrangian (and the principle of least action in classical physics). We cannot even say that the fundamental theory is equivalent to some Lagrangian approach in its effects (for example, we might say that Hamilton's or Newton's approach most properly describes physical law; and that we can express these in terms of an action is just accidental); but not so in reality – all we need is to be to show that the Lagrangian approach is mathematically equivalent to the more 'fundamental' theory, and we can use the same equivalence to send the results of Noether's theorem back into the other representation of physics. Given the overwhelming evidence supporting the current physical models, it would take a brave philosopher indeed to dispute them. Of course, Hume cannot be blamed for not realising this these developments came after his time. But modern philosophers who take Hume's argument seriously ought to take account of physics.

So the account is that numerous things are preserved throughout each interaction between a cause and an effect; and if a rabbit plopped into or out of existence then it would violate these conservation laws. We may imagine it, but by mathematical law it cannot be. There must therefore be something proceeding the rabbit, giving it its energy and momentum, and which loses that energy and momentum after the rabbit comes into existence; and this is nothing more than the relationship between the efficient cause and the effect. Hume missed from his list 'resemblance, proportions in quantity and number, degrees of any quality, and contrariety' mathematical and geometrical considerations concerning symmetry.

In any case, Hume continues,

> Since it is not from knowledge or any scientific reasoning, that we derive the opinion of the necessity of a cause to every new production, that opinion must necessarily arise from observation and experience. The next

> question, then, should naturally be, how experience gives rise to such a principle? But as I find it will be more convenient to sink this question in the following, Why we conclude, that such particular causes must necessarily have such particular effects, and why we form an inference from one to another? we shall make that the subject of our future inquiry. It will, perhaps, be found in the end, that the same answer will serve for both questions[15].

Now here again, we must remind ourselves of the Aristotelian answer to this question: that every substance has in itself, as part of its definition, various final causes, which describe what its various possible effects could be and what could not be a possible effect. Hume, of course, is committed to the founding principle of the enlightenment: that there are no final causes.

Hume once again reminds us of his principle that chains of knowledge must start and end with observation; we read or hear the words of a historian at one end of the chain, whose account ultimately goes back to various eyewitnesses. Similarly all reasoning concerning cause and effect must begin with an observation; even if all we recall is that the conviction that observation produced (memory in itself being flawed). Thus we need to trace our thoughts from the initial impression, to the idea of cause and effect, and from that to the nature of the idea.

> It is easy to observe, that ... the inference we draw from cause to effect, is not derived merely from a survey of these particular objects, and from such a penetration into their essences as may discover the dependence of the one upon the other. There is no object, which implies the existence of any other if we consider these objects in themselves, and never look beyond the ideas which we form of them. Such an inference would amount to knowledge, and would imply the absolute contradiction and impossibility of conceiving any thing different. But as all distinct ideas are separable, it is evident there can be no impossibility of that kind. When we pass from a present impression to the idea of any object, we might possibly have separated the idea from the impression, and have substituted any other idea in its room[16].

Again, here Hume directly denies the Aristotelian account. His argument, however, is weak: for each object we form a distinct idea; each distinct idea is separable; therefore the objects are separable and do not depend on each other. But here, Hume is begging the question; the very issue under debate is whether we can consider objects as separable entities in every respect. The idea is not the form of the object, but that which we use to grasp the form of the object; and if the form is necessarily connected to the existence of other objects (namely its final and efficient causes), then we cannot fully consider an object by itself without considering where it came from and what it might become.

Now we come to the central part of Hume's argument. We only infer causes by observation of regularities. We observe that something frequently follows another, with the properties of priority in time and conjunction in space, and by observing the repeated pattern we come to view this as a matter of cause and effect. Causality is therefore nothing but observed conjunction and correlation. We observe a constant conjunction between a flame and heat, and therefore assume that the flame causes the heat. This constant conjunction is, according to Hume, a third property of cause and effect. This does not help Hume; further observations only repeat what is already known, and cannot lead to the underlying reason for the relation between cause and effect. Next Hume asks whether experience produces

[15]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section III.

[16]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section VI.

the idea of causality by means of the understanding or intellect. Hume casts over the reasons we might suppose that the idea of causality emerges from the intellect. If it is due to the intellect, we might suppose that instances we do not experience would resemble those we do experience, and nature continues uniformly the same. But how can we reason that those events which we have had no experience will resemble those of which we have had experience? Hume answers that the only answer that we can refer to is from probability; that we have seen it numerous times, it is immensely probable that it will occur again. This probability, of course, is not intrinsic to materials: it cannot be calculated from first principles, but only comes from observation.

Let me use the example of a dice, and we wish to know what the probability is that it will land on a certain side. There are two approaches we can take. The first is to roll it numerous times, and observe the results. This is the observed "probability" (it would be better to call it an observed frequency). The second is to measure the distribution of the weight; consider the physical laws, and deduce the probability. We need not roll the dice once. This I call an intrinsic or calculated probability. Now these two methods should give the same results within errors. Hume only considers the first type of probability; indeed, according to Hume, the second type depends on the first, for the physical law we used to calculate the probability is (according to Hume) derived from numerous examples of measurement. Again, this may be questioned, for in modern physics physical law is more than just the result of repeated observation, but is merely tested by observation. There are many other principles entering our formation of the standard model of particle physics which though confirmed by observation are not derived from observation, such as symmetry and renormalisability. But for Hume, probability, coming from our observation of the conjunction of cause and effect, cannot be used to explain the relation that is cause and effect.

So what are we left with? Hume dismisses the idea of an intrinsic power (or final cause) – since he has already dismissed all the reasons that we might believe that to be the case.

> It shall therefore be allowed for a moment, that the production of one object by another in any one instance implies a power; and that this power is connected with its effect. But it having been already proved, that the power lies not in the sensible qualities of the cause; and there being nothing but the sensible qualities present to us; I ask, why in other instances you presume that the same power still exists, merely upon the appearance of these qualities? Your appeal to past experience decides nothing in the present case; and at the utmost can only prove, that that very object, which produced any other, was at that very instant endowed with such a power; but can never prove, that the same power must continue in the same object or collection of sensible qualities; much less, that a like power is always conjoined with like sensible qualities. Should it be said, that we have experience, that the same power continues united with the same object, and that like objects are endowed with like powers, I would renew my question, why from this experience we form any conclusion beyond those past instances, of which we have had experience. If you answer this question in the same manner as the preceding, your answer gives still occasion to a new question of the same kind, even in infinitum; which clearly proves, that the foregoing reasoning had no just foundation[17].

Hume concludes that reason can never show us the connection between cause and effect; therefore it must be simply be a matter for the imagination. There is no real discernible reason why one effect follows the cause, just the conjunction of our ideas.

---

[17]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section VI.

> Thus not only our reason fails us in the discovery of the ultimate connexion of causes and effects, but even after experience has informed us of their constant conjunction, it is impossible for us to satisfy ourselves by our reason, why we should extend that experience beyond those particular instances, which have fallen under our observation. We suppose, but are never able to prove, that there must be a resemblance betwixt those objects, of which we have had experience, and those which lie beyond the reach of our discovery.
>
> We have already taken notice of certain relations, which make us pass from one object to another, even though there be no reason to determine us to that transition; and this we may establish for a general rule, that wherever the mind constantly and uniformly makes a transition without any reason, it is influenced by these relations. Now this is exactly the present case. Reason can never show us the connexion of one object with another, though aided by experience, and the observation of their constant conjunction in all past instances. When the mind, therefore, passes from the idea or impression of one object to the idea or belief of another, it is not determined by reason, but by certain principles, which associate together the ideas of these objects, and unite them in the imagination. Had ideas no more union in the fancy than objects seem to have to the understanding, we could never draw any inference from causes to effects, nor repose belief in any matter of fact. The inference, therefore, depends solely on the union of ideas.
>
> ...
>
> But though I acknowledge this to be a true principle of association among ideas, I assert it to be the very same with that between the ideas of cause and effect, and to be an essential part in all our reasonings from that relation. We have no other notion of cause and effect, but that of certain objects, which have been always conjoined together, and which in all past instances have been found inseparable. We cannot penetrate into the reason of the conjunction. We only observe the thing itself, and always find that from the constant conjunction the objects acquire an union in the imagination. When the impression of one becomes present to us, we immediately form an idea of its usual attendant; and consequently we may establish this as one part of the definition of an opinion or belief, that it is an idea related to or associated with a present impression.
>
> Thus though causation be a philosophical relation, as implying contiguity, succession, and constant conjunction, yet it is only so far as it is a natural relation, and produces an union among our ideas, that we are able to reason upon it, or draw any inference from it[18].

So for Hume causation is merely a repeated pattern of regularity; it is not inherent within the substances (or we have no reason to believe that it is inherent within the substances), but merely something that usually happens. We create the connection of causality in our mind because of a constant conjunction, but there is no such connection in nature, or if there is, we have no good reason to be sure of its existence. When we strike a match, there is nothing about the match itself that causes it to light up after the application of a small amount of heat through the work of friction; we just observe that when we strike a match it frequently leads to a flame and for this reason we create a connection as though match + friction = fire. In effect, this is a removal of the Aristotelian efficient causality in addition to the final and formal causality that Hume dismisses. Hume has considerably more to say

[18]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section VI.

on the subject of causality, but to be brief before commenting on this picture, I will just summarise two further aspects of his discussion which will bear on material I need later.

The first is Hume's discussion of probability. Hume sees chance as contrary to causation. Since he believes causation to be a necessary connection, this is obvious. In practice, of course, there is (according to Hume, and classical physics, though not the correct quantum theory, agrees with him) no chance in reality, merely ignorance of the underlying effects. Chance being the negation of a cause, we cannot rank one chance as superior to the other. To believe otherwise must be to believe that there is some cause that makes red superior to blue, but this contradicts the notion that chance is a negation of cause. Take, for example, the case of a die that has four red sides and two blue sides. We might suppose that this makes it twice as likely that it lands on blue than it does red, but Hume disagrees. We cannot decide this by mere examination of ideas, and it is impossible to prove that blue will dominate (for example the dice might be loaded). Probability, according to Hume, can thus only be observed by measurement; by rolling the dice and counting how many times it lands as red and how many times it lands as blue. As we do so repeatedly, we gradually gain confirmation of our first intuition, that as there are twice as many blue sides as there are red it is twice as likely to land blue as it is red. But this picture cannot be determined from first principle; it is a picture that gradually emerges through repeated observation. So it is when we consider the chances of causes. With few observations, we are uncertain about the ideas of causes, but with more observations, we gradually form a clearer picture.

> As the habit, which produces the association, arises from the frequent conjunction of objects, it must arrive at its perfection by degrees, and must acquire new force from each instance, that falls under our observation. The first instance has little or no force: The second makes some addition to it: The third becomes still more sensible; and it is by these slow steps, that our judgement arrives at a full assurance. But before it attains this pitch of perfection, it passes through several inferior degrees, and in all of them is only to be esteemed a presumption or probability. The gradation, therefore, from probabilities to proofs is in many cases insensible; and the difference betwixt these kinds of evidence is more easily perceived in the remote degrees, than in the near and contiguous. ...
>
> It is evident, that when an object is attended with contrary effects, we judge of them only by our past experience, and always consider those as possible, which we have observed to follow from it. And as past experience regulates our judgement concerning the possibility of these effects, so it does that concerning their probability; and that effect, which has been the most common, we always esteem the most likely. ...
>
> There is no probability so great as not to allow of a contrary possibility; because otherwise it would cease to be a probability, and would become a certainty. That probability of causes, which is most extensive, and which we at present examine, depends on a contrariety of experiments; and it is evident an experiment in the past proves at least a possibility for the future. ...
>
> All our reasonings concerning the probability of causes are founded on the transferring of past to future. The transferring of any past experiment to the future is sufficient to give us a view of the object; whether that experiment be single, or combined with others of the same kind; whether it be entire, or opposed by others of a contrary kind. ...
>
> It is obvious, that in reasonings of this kind, it is not the object presented to us, which, considered in itself, affords us any reason to draw a conclusion concerning any other object or event. For as this latter object is supposed uncertain, and as the uncertainty is derived from a concealed

> contrariety of causes in the former, were any of the causes placed in the known qualities of that object, they would no longer be concealed, nor would our conclusion be uncertain.
>
> But, secondly, it is equally obvious in this species of reasoning, that if the transference of the past to the future were founded merely on a conclusion of the understanding, it could never occasion any belief or assurance. When we transfer contrary experiments to the future, we can only repeat these contrary experiments with their particular proportions; which could not produce assurance in any single event, upon which we reason, unless the fancy melted together all those images that concur, and extracted from them one single idea or image, which is intense and lively in proportion to the number of experiments from which it is derived, and their superiority above their antagonists. Our past experience presents no determinate object; and as our belief, however faint, fixes itself on a determinate object, it is evident that the belief arises not merely from the transference of past to future, but from some operation of the fancy conjoined with it. This may lead us to conceive the manner, in which that faculty enters into all our reasonings[19].

I should also observe that Hume also recognises probability through analogy; we see an object which is similar to one we have already studied and might reasonably be supposed to have the same chances.

Hume thus separates what we may call mathematical probability from practical probability. Mathematical probability is the ability to pre-calculate the chances of events occurring in some ideal situation; practical probability is the observation of past events combined with the assumption that future events will follow the same pattern. Hume claim is essentially that mathematical probability (derived from the object considered in itself) is impossible to apply to the real world; that only practical probability is useful beyond the abstract (and useless) study of mathematics. Equally I should emphasise that for Hume, there is no certainty, only probabilities that approach one[20].

Finally, I wish to discuss Hume's distrust of necessity. Recall that he described causality as a necessary connection. However, since the connection is only one we make in our minds by observation of past conjunctions, and observation cannot bring us certainty but only increase the probability, the idea of there being a necessary connection is impossible.

But if there is no necessary connection between cause and effect, what shall we conclude? I see no alternative, if Hume's premises and arguments are correct, but to agree with him when he says

> According to the precedent doctrine, there are no objects, which by the mere survey, without consulting experience, we can determine to be the causes of any other; and no objects, which we can certainly determine in the same manner not to be the causes. Any thing may produce any thing. Creation, annihilation, motion, reason, volition; all these may arise from one another, or from any other object we can imagine. Nor will this appear strange, if we compare two principles explained above, that the constant conjunction of objects determines their causation, and that properly speaking, no objects are contrary to each other, but existence and non-existence. Where objects are not contrary, nothing hinders them from having that constant conjunction, on which the relation of cause and effect totally depends[21].

---

[19]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section XII.

[20]I should comment here that I disagree with Hume's distinction between probabilities and certainties. A certainty is a probability of one; it belongs to the same class of object.

[21]Hume, *A Treatise on Human Understanding*, Book 1, Part III, section XV.

8.5.2. Response to Hume

So what shall we make of this? I have discussed Humean causality in detail, because I see it as the heart of many objections to the arguments that I shall give later, and because of Hume's historical importance. If Darwin was the father of modern atheism, the man who made it intellectually respectable; then Hume was perhaps one of its grandfathers, providing the underlying metaphysical basis that many atheists would adopt, and supplying many of their strongest arguments against theism. Before I can make the positive arguments, it is necessary that I dismiss Hume and thus the objections based on his work, and ultimately on his premises. One would have thought that Hume's final conclusion is in itself enough to show that he must have made a mistake. Hume states that anything can give rise to anything else. A quark can decay into a bunch of flowers. Even in classical mechanics, this is a gross violation of the principles of energy and momentum; which in turn are derived from fundamental symmetry principles.

Let us consider the case of the solar system. Kepler accurately (though not perfectly accurately, but for the first part of this discussion I will consider the thought experiment of a solar system described by Newtonian gravity) described the motion of the planets around the sun. It is important to note that Kepler's laws are as far as we can go through observation alone. This is what the constant conjunction of observation tells us: that the planets move around the sun in perfect ellipses with the sun as the focus of one of the ellipses. According to Hume, we observed this yesterday, and last week, and the week before that, so we have good confidence that they will continue to move in the same way tomorrow. Newton, on the other hand, noted that this is the behaviour that we would expect if his laws of motion were obeyed and the sun exerted a force on the planets proportional to the product of their masses and inversely proportional to the square of the distance between them. Extending this to all bodies, we can also explain the motion of the moon around the earth and our attraction towards the earth. In other words, the observed motion of matter is exactly what we would expect as though each body produced a gravitational field as described by Newton. But Hume denies that we can convert this 'as though' into a 'because'. We do not see the field; merely the movement of the planets.

This is not enough; the physicist is not satisfied with merely 'as though': the situation cries out for a deeper explanation; and the most obvious explanation is that bodies do in practice generate a gravitational force field. Now, it may be objected that Newton's gravity is, in fact, wrong. It does not explain the orbit of mercury correctly; while Einstein's picture of gravity, that bodies move along geodesics of the metric of space time, and that the metric is bent by massive bodies according to Einstein's field equation does; and hopefully soon Einstein's theory will be supplemented by a quantum theory of gravity which makes the same prediction for the motion of the solar system. We should not then say that the explanation of Einstein is correct, or of some model of quantum gravity, but that there is an explanation which is either classical general relativity, or some quantum general relativity, or some other of the few classes of models which lead to the same effect; but is most certainly not Newtonian Gravity, nor any one of a much larger class of models which give incorrect predictions. The physicist therefore says that there is a mechanism for cause and effect, and it is either A or B or C (though we do not yet know which), but is certainly not α or β or γ[22]. Hume, on the other hand denies that can be such an explanation, or at the very least

[22]Two comments should be made here. Some will now want to use Ockham's razor to choose between A, B, or C; saying whichever explanation is the simplest should be preferred. My own feeling is that the use of Ockham's razor has done far more harm than good, and it should be avoided. There is no reason why we should expect the simplest reason to be correct, and the notion of 'simplest' itself is difficult to unambiguously define. We may, for example, define it as the explanation with the fewest independent assumptions; however, even identifying and counting assumptions and showing their independence is somewhat subjective since we have an incomplete knowledge. It is better just to leave all the options on the table; there is no shame in saying that the evidence is too slim to judge between them. Secondly, this seems to be in agreement with Popper's notion of falsification. It should be noted that it only refers to one small, though important,

that we can know anything about it beyond constant conjunction and that cause proceeds effect in time. Hume reduces science to just experiment and phenomenology; there is no place in his world for theory. He does not consider that we can make negative statements about the particular causes of some being, and by eliminating the alternative options come to a positive statement concerning causality.

I think that it was for this reason that Werner Heisenberg, one of the most important twentieth century physicists and a father of quantum mechanics, who wrote

> The three philosophers who can be taken as representatives for early empiricist philosophy are Locke, Berkeley and Hume. Locke holds, contrary to Descartes, that all knowledge is ultimately founded in experience. This experience may be sensation or perception of the operation of our own mind. Knowledge, so Locke states, is the perception of the agreement or disagreement of two ideas. The next step was taken by Berkeley. If actually all our knowledge is derived from perception, there is no meaning in the statement that the things really exist; because if the perception is given it cannot possibly make any difference whether the things exist or do not exist. Therefore, to be perceived is identical with existence. This line of argument then was extended to an extreme scepticism by Hume, who denied induction and causation and thereby arrived at a conclusion which if taken seriously would destroy the basis of all empirical science (Heisenberg, 1958).

So if Hume's conclusion is dangerous, and his reasoning solid, then the problem must lie in one of his premises. I will discuss four of his premises: firstly his definition, that causation is a necessary connection; secondly his proposition that chance negates causation; thirdly his presumption that our knowledge of the real world can only come through direct observation; and fourthly the enlightenment assumption that that objects cannot have a final cause, or if they do we cannot know about it. We cannot discover anything about the effect by considering the nature of the cause. Let me examine each of these premises in turn.

(1) **Causation is a necessary connection.** The key point that I disagree with here is *necessary*. An object need not have a single final cause. A particle (let's say a Higgs Boson) may decay in several different ways, perhaps into a $\tau\bar{\tau}$ pair or a W^+W^- pair. The word *necessary* carries the implication that the Higgs Boson must decay in one particular way; or that the $\tau\bar{\tau}$ must come from a Higgs Boson. No defender of Aristotelian causality would agree with this. The Aristotelian lists the possible decay channels of the Higgs as its final causes, and states that for each individual $\tau\bar{\tau}$ pair there must have been an efficient cause; for some $\tau\bar{\tau}$s this might have been a Higgs, for others it might have been a photon, for others something else; but for each individual pair it must have been one of these, and not any of the others. Perhaps then, Hume is referring to this second sense of *necessary*: for every effect there must have been a cause which has the effect as one of its final causes. But this is nothing more than saying that things cannot just plop into existence for no reason; and this was not the subject of Hume's discussion at this point of his argument, which was focused on the observed conjunction between two beings. Hume said nothing for or against the 'plop' theory of 'causation'; so it cannot be this version of necessity that he was querying. Hume's definition of causation as a *necessary* connection is thus a straw man. Causation is therefore necessary in the sense that every effect must have some cause, but is not necessarily necessary in the sense that every cause must lead to the same effect or every similar effect must have had the same type of cause.

---

aspect of science: that of choosing between competing expectations. Other aspects of science, such as much experimental work, are not based on falsification but on positivism and description.

Now it is true that Hume mostly focuses on the *connection* part of his definition of causation; but here again we may consider what it would be if there was no connection between cause and effect. We observe the cause ceasing to exist, and the effect coming into existence. If there is no connection between the two, then we are back to a version of the 'plop' theory: the cause ceases to exist for no reason, and the effect starts to exist for no reason and just happens to do so simultaneously with the cause. Such an assertion is surely unreasonable. Or perhaps Hume simply asserts that we cannot in principle know what the connection is; if so, he, speaking at the beginning of modern science, was being presumptuous about what we may know at the end. The various conservation laws of physics certainly give a ground for understanding the connection.

(2) **Chance negates causation:** This presupposes that causation is a necessary connection in the first sense discussed above, and I have shown that this need not be true. The decay channels of the Higgs Boson each occur with a certain, calculable, probability. There is both chance (in the sense that any of the decay channels can occur) and causation (in that once they occur, the Higgs ceases to exist and its energy and momentum is given to the products of its decay; there is an efficient cause and final cause). Chance is inconsistent with some formulations of causality, but not all, and in particular not Aristotle's substance causality.

(3) **Our knowledge of the real world can only come through direct observation:** This is the assumption of empiricism. Now, I would certainly stress the importance of empirical investigation in the formulation of physical law. We need it to inspire us; and we require it to choose between competing models (as well as giving the theorist inspiration when he is out of ideas). However, empiricism goes further than this, and states that theory can only depend on observation. This is false; we also use the premises of mathematical consistency, and, particularly in modern physics, various symmetry and topological principles in our construction of physical law. Not every conceivable theory satisfies these conditions. How far can this take us? It gives us the ability to predict, which the Humean conception of the laws of physics as merely the documentation of observed regularity cannot. For example, the standard model of particle physics is, as we will see, primarily built from requiring certain symmetries to hold; that it is a quantum field theory, and that it is mathematically consistent. To say that nature is described by a quantum rather than a deterministic theory comes from experiment; the types and numbers of each species of particle comes from experiment; but the symmetries from principle as much as anything else (though confirmed by experiment). These premises are each discrete quantities; they do not have error bounds. But if we add to these empirically derived premises the requirement for mathematical consistency, we have no freedom left in the construction of the model beyond a finite number of non-fixed parameters which we have to measure. There are many conceivable ways in which we could through together these ingredients, but only one which is mathematically consistent (again, up to the various parameters)[23].

Now from the standard model, we had a prediction, not simply an analogy with similar particles. It's mathematical consistency required one more particle, the Higgs Boson, which has only recently been observed. Now think about this from a Humean perspective. Hume reduces physics to the observation of regularities; and, until we reached high enough energies, there was no indication in any previous experiment that the Higgs was there. We cannot say that it would not be there, but there was no reason to say that it would be there. Furthermore, we not

[23]Do not think from this that I am claiming that the standard model is the full answer: it excludes the symmetry that leads to general relativity. The remaining problem facing particle physicists is how to construct a model which is mathematically consistent and which incorporates this symmetry.

only predicted the existence of the Higgs Boson, but calculated its decay channels and the probability of its various decays. Compare this to Hume's discussion of probability; that given the assumption of empiricism we can only construct probabilities on the basis of past observation. We calculated it in advance by considering the nature of the Higgs, something which Hume denied was possible, and are now testing the calculation against experiment. This test may of course fail; if so, then the standard model needs to be replaced by some other model which extends it (and many such models have been proposed; indeed, we know that the standard model is not the full answer); but Hume's argument was that it was impossible in principle to do so, and this is clearly false. If the test succeeds, then the discovery of the Higgs Boson will stand as the greatest vindication of twentieth century physics.

The standard model provides us with an explanation of the connection between cause and effect which Hume could not observe and thus denied. Because it was constructed through more than just empirical input, it avoids the difficulties which Hume highlighted. From the perspective of pre-Newtonian physics, Hume's explanation for causality is perfectly reasonable; indeed, perhaps even the best way of thinking about the subject. But the progress of Newtonian and post-Newtonian physics has shown Hume to be wrong. Hume himself should not be criticised for this; the mathematical structure, the understanding of the importance of symmetry all came after him; indeed, the final pieces of the puzzle only arrived in the latter part of the twentieth century. Does this mean that empiricism, or at least Hume's version of it, should therefore be abandoned as untenable given our current knowledge and capabilities? I believe so, although demonstrating this would require considerably more argumentation than I can provide here. However, modern defenders of Hume need to be aware of the criticisms of his approach by modern physicists, and to attempt to answer them. But we should note that with his empiricism, most of Hume's thought also goes, and much of the thought of his later disciples.

(4) **Objects cannot have a final cause; or we cannot know about it:** This is, of course, one of the founding assumptions of enlightenment philosophy, and we noticed how it played out in Hume's thought: denial of the final cause entailed denial of efficient causes and the destruction of modern physics. Hume's argument against final causality is, however, based on his empiricism, which we have seen is false. Instead, as we shall see presently, modern physics not only allows for final causality, but allows us to state the principle more precisely than Aristotle or his followers were able to do.

I should, however, be fair to Hume: my dismissal of him is anachronistic. The 'laws' available in Hume's day: Kepler's laws of motion; Boyle's laws of gases, Galileo's dynamics, the various laws and principles of optics or magnetism, for example, were just collections of experimental results. In that time, it was quite reasonable to suppose that that was all that science could achieve; that all physical law would take that form. Hume was the opposite of Descartes (or the popular caricature of Descartes) who held that reason alone was reliable; but Hume's opposite view that observation alone is reliable turned out to be equally flawed: it is Galileo's compromise position that proved best. The process of formalisation of physical law and the awareness of how these physically derived 'laws' could be derived from theoretical – what Hume would have called metaphysical – principles started with Newton, though Newton perhaps did not realise the full consequences of what he was beginning. Thus Newton himself showed how Kepler's and Galileo's works followed as consequences his principles of classical mechanics; it would take until the days of Maxwell and Boltzmann for optics and the theory of gases to fall into place as pieces of Newton's system, and, of course, today we see that Newton's 'laws' are themselves merely consequences

of relativistic quantum mechanics applied to a certain regime, moving further away from the initial experiment, and deeper into an unobservable but nonetheless powerful formalism. Hume should thus be thought of as a failed prophet: he saw the signs of his times, read into them a set of ideas which with the knowledge available to him and in his day were quite reasonable, and got it completely wrong. The problem is not with Hume, but with his modern day devotees who, understanding how science has developed, should know better.

8.6. Is Enlightenment a Good Thing?

DEFINITION 8.6.1. **Enlightenment**:
I define an enlightenment philosophy as one which denies the existence, or at the very least usefulness, of the final and formal causes. △

Objection:
This is not how the early enlightenment philosophers viewed themselves. They claimed that the enlightenment is the triumph of reason over superstition; the victory of empirical investigation over relying on the opinions of ancient authorities, and most of all the promotion of intellectual freedom above any previous dogma. No truth should be imposed on subsequent generations; everything must be open to question. To oppose the enlightenment is to oppose reason and empirical investigation, the twin foundations of modern science.

To take a specific example, Kant, in his famous essay, wrote,

> Enlightenment is man's emergence from his self-imposed nonage. Nonage is the inability to use one's own understanding without another's guidance. This nonage is self-imposed if its cause lies not in lack of understanding but in indecision and lack of courage to use one's own mind without another's guidance. Dare to know! (Sapere aude.) "Have the courage to use your own understanding," is therefore the motto of the enlightenment. ...
>
> This enlightenment requires nothing but freedom–and the most innocent of all that may be called "freedom": freedom to make public use of one's reason in all matters. Now I hear the cry from all sides: "Do not argue!" The officer says: "Do not argue–drill!" The tax collector: "Do not argue–pay!" The pastor: "Do not argue–believe!" Only one ruler in the world says: "Argue as much as you please, but obey!" We find restrictions on freedom everywhere. But which restriction is harmful to enlightenment? Which restriction is innocent, and which advances enlightenment? I reply: the public use of one's reason must be free at all times, and this alone can bring enlightenment to mankind. (Kant, 1784)

•

Response:
There is nothing in this self description which is all of unique to the enlightenment (i.e. was not held by earlier scholars), fully held by enlightenment scholars, and beneficial. What was new in the enlightenment was not good, and what was good in it was not new. Take, for example, the claim that the enlightenment is the bastion of reason. Was there any group of men more devoted to logic and reason than the medieval scholastics? And certainly the enlightenment philosophers have not all been consistently rational; even the best of them contain many fallacies. After all, there is the school of Hume, of Descartes, and Kant, and

many others all contradicting each other in many respects. At most one of them can be right; the others must have made mistakes.

Is the enlightenment to be defined in terms of an opposition to superstition? If this is to be defined as the assumption that no supernatural occurrences are possible, then this assumption merely limits our intellectual freedoms, and perhaps disastrously. If superstition is to reject any supernatural occurrences which lack good evidence, then the churchmen whom Kant opposed would have been the first to agree with him. Nor can empirical investigation be the cornerstone of enlightenment, since Aristotle based his philosophy on his own observations of nature, as did Ptomely. Indeed, Galileo, that great hero of enlightenment, was rejected not because he challenged authority but because Copernicus' system, which he championed, contradicted the best observations of the time – something he could not refute. The ancient Chinese built their system on empirical investigation and tested it as thoroughly as they could — that doesn't mean that they were enlightened (in Kant's sense of the word, which should be distinguished from what the Chinese, influenced by Buddhism, would have meant by the term). Of course, we moderns perform experiments better than any before us; we are more careful with the observations, more systematic, have better understanding of the necessity of controls, and most importantly are far more careful in documenting and accounting for systematic and other experimental errors. But that is not a matter of a change of principle, but merely improvement in method. One does not have to be enlightened to appreciate it.

Was the enlightenment principally about freedom of thought? But, of course, our thought should not be free: we should be constrained by the truth. If something contradicts our observations, then it must be rejected, and to continue to insist on it, or accept those who do as having a valid opinion, is the height of folly. Nor should we accept any view which is inherently self-contradictory. I do not think that any supposedly enlightened scholar would disagree with this. It is this sort of restriction that the churchmen believed themselves to be imposing, and which Kant objected to in his essay. Christianity is, of course, a revealed religion: based ultimately on observation rather than theory. Its dogmas are the collection together of those observations, although, of course, many of them can also be confirmed through well-reasoned philosophy. One can question, of course, whether those observations are valid, and many do; but if we accept them, then obviously our reason has to be constrained by them. What we might think of as absolute freedom is, of course, no freedom at all. Freedom requires boundaries. If we were truly free to reject all that went before us, and each generation started out by doing the same, then we would not make any progress at all: there would just be intellectual chaos.

Equally, the rejection of previous views is clear throughout classical philosophy. Socrates, Plato, Aristotle; all of these displayed immense original thought. In the medieval period, the conversion from Neo-Platonism to Aristotelianism (despite the difficulties of a naive interpretation of Aristotle) was in immense change, which displayed precisely the sort of freedom of thought that Kant advocated. The best progress has always been made when we keep one foot in the old and the other reaching out to the new; and this is precisely how the medieval scholars operated. By cutting themselves off from the old, the enlightenment philosophers opened themselves into falling into the subtle mistakes that their predecessors had warned against, and many of them did fall into those mistakes. Instead, I have found few thinkers less free than those who have uncritically accepted the dogma of the enlightenment philosophers.

Was the enlightenment about the rejection of dogma? Not if Kant's essay is any guide. Kant argues that we should be free to consider things apart from dogma. But what if our scholar should conclude that the best reason leads him to accept that the dogma was, in fact, correct after all? Should we forbid him from coming to that conclusion? Certainly not!

Is enlightenment to be identified through a sceptical approach to Philosophy? I do not think that it can; there were sceptics before Descartes (such as Al-Ghazali and Sanches), and

not every Philosopher after Descartes but in the general enlightened tradition was a sceptic in the complete sense of the word. Almost everyone is sceptical to a certain degree: after all, the Christians were sceptical of the arguments put forward by the pagans, the Jews sceptical of the Muslims, and the Buddhists would be sceptical about all of them. But surely these thinkers were overly credulous when it came to those of their own tradition? Not entirely, as the various theological disputes in each tradition have shown. Few Christian thinkers of any note (though doubtless there are many exceptions among the numerous Christian unthinkers) once they reached maturity accepted the truth of the resurrection on authority, but because they believed it to be the best explanation of the evidence. But to examine the evidence thoroughly means questioning it, even if at the end of the day you conclude that your teachers were right all along. But if many Christians (for example) were too slow to question those of their preferred tradition, can we not say the same of many modern scholars in the enlightenment tradition? How many have just followed Hume or Kant or Wittgenstein without deeply questioning if they were right? For example, Hume's criticisms of scholasticism almost always, and obviously to anyone who has given both Hume and the medieval thinkers more than a superficial reading, completely missed the mark, attacking straw men rather than the arguments they actually put forward. Yet far too many modern thinkers just cite Hume as having disposed of that thought without demonstrating that he did. Given that many non-enlightenment philosophers employed scepticism, and that the principle was not employed to any greater degree than is customary by many enlightenment philosophers, we cannot use this as the distinguishing mark of enlightenment.

So we come down to the question of authority. Of course, some areas of study would be impossible without relying on the witness of ancient authorities; history being the most obvious example. We would have to weigh between different sometimes contradictory accounts, but this is no different from the philosopher when choosing which of his predecessors to agree with and which to reject. Even as a physicist, I rely on authority. I have not performed every experiment which shapes my views; some of them would be impossible for me to do alone. Instead, I have relied on the trustworthiness of those who performed those experiments. Indeed, I have mostly accepted those accounts second or third hand through my teachers or the textbooks I used to learn the subject. It is always good to check something for yourself (and we must as much as we can if our own work will depend on it); it is not possible to check everything. But disputing authorities is again not unique to the enlightenment. Thomas Aquinas himself argued that the argument from authority was the weakest of them all (Aquinas, 1274). By accepting Aristotle, he rejected Plato, Parmenides, Cicero and the rest: why? Because he believed that Aristotle gave the most reasonable account of the world in classical philosophy. But he did not accept all of Aristotle: he rejected those parts of Aristotle which seemed to him to be less good than others. He took from the past that which was good, and left behind that which was bad, and it was his reason that allowed him to distinguish between the two. Why then did Aquinas cite so many past authorities? Because he recognised that if you disagreed with a great thinker from the past, you needed to be able to show where they went wrong; if they agreed with you, then that added confidence that your argument was correct. However, Aquinas' main thrust in his major works was always based on an argument from first principles. He cited authorities, but nowhere relied on them.

One does not rationally blindly accept authorities; nowhere have the leading scholars of the church asked people to do so. One should first investigate whether the authority is reliable. Suppose that we conclude that the authority is reliable (at least up to a certain extent, or in a certain area of competence). Is it then rational to reject that authority, to say that it is wrong simply because it is an authority? By no means! If an authority is shown to be trustworthy, then it is right to trust it, at least as far as it is competent. To blindly reject authorities is just as bad as blindly accepting them. We accept or reject them based on how consistent they are with everything else, and how trustworthy we deem them to be. If we

can check their conclusions directly, follow or reproduce their arguments, then it is best to do so. But checking them does not mean that we reject them. If we find a contradiction between our conclusions and theirs, then our first instinct should be to assume that we made the mistake. Only after a careful checking of both arguments should we conclude that we are right and they wrong. For example, one does not join the Eastern Orthodox church and then forcibly submit one's reason to follow the ecumenical councils. Instead one uses ones reason to test the findings of the ecumenical councils, and, after finding them reliable, that becomes part of the reason for accepting Eastern Orthodoxy.

Equally the modern philosophers and their disciples also rely on authorities – just different authorities. How many times, for example, have I read that Kant and Hume destroyed the classical arguments for God, without any discussion about whether this was truly so? How many times have I seen definitions drawn from Kant without any questioning of whether those definitions are useful? How many times have people accepted the idea that medieval philosophy was disproved on the authority of the renaissance or early modern scholars, without surveying the evidence to see if it were truly so?

But, of course, one can not reject authorities as such. One does not, as a purveyor of ideas, reject Plato, Aristotle, Aquinas, Hume Descartes, or Kant in themselves. One rejects either their conclusions, and therefore either their arguments, or their premises (which may in turn be based on observed evidence). When Kant tries to argue that one must be free to reject previous dogma, his free thinker who disagrees with that dogma, as long as he is rational, must either be saying that that dogma was entirely arbitrary (which was obviously not the case, as any student of Medieval or early Christianity would be able to testify), or he rejects the premises on which that dogma was based. So we are then led to the conclusion that the only difference between the best of the enlightenment philosophers and the best of those whom they rebelled against (the worst in both traditions, of course, merely followed the crowd, whether that crowd was led by Aristotle on one hand or Kant on the other – with the enlightened group simultaneously patting themselves on the back for their originality and rationality) was that there were some premises accepted by the scholastics which were universally rejected by the enlightenment philosophers. So what were those premises? I have offered up the acceptance of formal and final causality as one possibility. Certainly this was agreed on by the scholastics I have studied the most, and rejected by the principal founders of the enlightenment tradition. Perhaps a true historian of philosophy could provide a better suggestion, but for the purposes of this work the definition of Enlightenment I have given above is reasonable enough to be used.

So I certainly do not accept that rejecting enlightenment philosophy means rejecting the methods and findings of modern science. On the contrary, I firmly believe (though obviously have not proved in the discussion so far) that to accept modern science means that we should reject supposedly enlightened philosophy. But one thing that all modern philosophers have agreed on is that formal and final causes are either irrelevant or non-existent. If science agrees with this assessment, then modern philosophy might be correct. If it disagrees with it, and form and finality still have a place, then the various schools of enlightenment philosophy should be rejected. o

8.7. The Fate of Scholastic Metaphysics

The pre-eminence of Newton in physics and mathematics, and his own success and the even greater success of those who followed him, also led to the unconscious absorption of most of his metaphysical presuppositions; to the extent that even today many people would not even consider that they could be reasonably challenged. The net result of this is that Scholastic Metaphysics is seen in many quarters as something we are well rid of. In particular, there is a particular dislike of teleology. But it is worth reviewing the reasons why scholastic metaphysics was rejected. These reasons included:

(1) The terms of the metaphysicists were seen as imprecise and ill-defined, particularly in comparison to the new mathematical physics.
(2) There was a lack of progress in physics by the scholastics; and in particular they appealed to dogma and authority ahead of experiment and innovation.
(3) Aristotelian physics got some things badly wrong, as did Ptolemy's astronomy.
(4) The new mechanical metaphysics led to enormous and rapid progress in physics, with Newton as the climax of this process, and (later) Maxwell and Einstein providing its culmination.
(5) The new mechanical physics makes no use of causes, and final causes especially, so these are superfluous and can be discarded.
(6) Searching for final causes led many people down blind alleys.
(7) Physics is best performed without any metaphysical presuppositions.
(8) Certain aspects of classical physics cannot be made to conform with scholastic metaphysics.

Before assessing these, it is worth reiterating two points. Firstly, that there was a general confusion between design and teleology in the early modern writers, which continued through their successors until the present day. The arguments over 'final causality' solely concerned themselves with design, and the idea that beings have tendencies towards only certain ends was not mentioned. The problem with this is that given efficient causality or even event causality, one has to have either a complete free-for-all, with anything possible, or some concept of final causality. The first of these was proposed by Hume, but can be rejected on account of the success of science. Thus final causality cannot be dismissed from physics; the best that can be done is that it is viewed as equivalent to efficient or event causality, and thus the concept is superfluous. The second point is that few[24] of the early modern thinkers attacked the rational basis of scholastic metaphysics, and those who tried failed to understand it, so their attacks were worthless. In particular, they did not address the issue of potentia and accidental change; which proved disastrous for much of the philosophy which followed Descartes. They did not show that scholastic metaphysics was self-inconsistent, nor address its basic premises, nor even that it contradicts modern physics. It was stated that it could not be understood, but this is clearly false, since many people did understand it and do to this day.

So let us look at the objections listed above. Firstly, there is the sense that key quantities were ill-defined. I think a more accurate statement would be that they were too broadly defined, covering a multitude of different possibilities. Now, it is clear to see why the scholastics had to do this: they did not have a precise understanding of physics. They thus needed some rather fuzzy definitions to cover all possibilities. Indeed, it could be said to be imprecise but accurate is better than to be precise and inaccurate. In any case, the solution to this problem is not to abandon the metaphysics, but to define the terms better, and in particular create definitions which are consistent with both the old philosophy and the new mathematical physics. So this is not an argument against the metaphysics, or a reason to abandon it, merely a request that it be done better.

The second objection is the lack of progress in science of the scholastics. One of the arguments for this is that saying 'That just happens to be its final cause' is a barrier to further inquiry. But is it? Scholastic metaphysics demands a connection between the four causes, and investigation of these connections would lead to further inquiry. It is not enough to say 'That is its final cause', one has to ask what it is about the form and the matter that implies that particular final cause. So scholasticism need not suppress inquiry. Then there is the historical question: is it true that there was no real progress in physics between Aristotle and Galileo? However, as we have seen, this statement is partially misleading and partially

[24] I say 'few' to be conservative; I have not encountered any serious and valid objections of this sort in the early modern literature. However, I have not read everything that was written in the period, so I am reluctant to say 'none'.

false. Scholasticism did not emerge until the texts of Aristotle reached Western Europe, were translated into Latin, and accepted by the intellectual community, a process that was not completed until the twelfth century. Before then, intellectual life was either based on Plato, or no philosophy at all, or, going back into the classical period, on the various competing schools of Greek Philosophy. True, Aristotle was also revered a little earlier in the Islamic world, but the wider culture of that context was opposed to any philosophy, whether proto-scholastic or proto-mechanical. So the idea that the Aristotelians had two thousand years to get their act together is false; they had about four hundred years (the last two hundred of which were eaten by the aftermath of the black death and then the Renaissance). But the second point is that they were not quiet during this time, since the new mathematical physics emerged in the thirteenth century. Certainly, progress was slow until the time of Galileo, but to say that progress in physics was absent in this period is simply false. True, there were two strands emerging in the universities at this time, with the theological philosophers still accepting Aristotle (though in divided camps), and the mathematical philosophers working on the new physics. It happened to be mechanists rather than scholastics who made the key breakthrough. But this doesn't disprove scholasticism. Just because modern science originated with mechanism does not prove mechanism to be true; just because the scholastic scholars of the period failed to realise the potential of the new mathematical physics and incorporate it into a minimally modified form of their system does not prove the scholastic philosophy false. The question is not how something originated, but whether or not it is true; good fruit can sometimes come from a bad seed. So this is not a disproof of the scholastic metaphysics, though it does cast doubt on it. The scholastics needed a Descartes or Galileo of their own, to merge the new physics with their philosophy, and they did not have one. This historical fact needs explanation. It might be a problem with the philosophy; or it might be a comment on the low quality of the scholastic scholars of the period; or it might be that nobody noticed that scholasticism needed modification and was prepared to attempt it; or some other reason. But this historical accident is not a refutation of the scholastic philosophy as a whole.

The third objection is that Aristotelian physics got some things badly wrong; which is certainly true. This means that Aristotle's physics is wrong. However, the question concerns Aristotle's metaphysics. To reject Aristotle's metaphysics on the basis of his physics being wrong is only a valid argument if Aristotle's metaphysics necessarily implies his physics (since the metaphysics is ontologically prior to the physics). But this is false, because Aristotle's physics is a combination of his metaphysics and some other assumptions concerning the nature of matter drawn from earlier Greek writers (such as the theory of the four elements, and so on). At least one of those two must be flawed for the physics to be flawed. However, it only needs one, and we know that the assumptions concerning the nature of matter were false. Thus the failure of Aristotle's physics does not imply the failure of Aristotle's metaphysics; particularly since the metaphysics is based on general considerations on the nature of change and constancy which must apply to any physical system. Thus this objection to Aristotle's metaphysics also fails to refute it.

Equally, this argument can be turned around. Part of the attack on formal and final causes was that physics can just be reduced to matter in locomotion. But there is more to physics than just locomotion. Spin (referring to the quantity in particle physics alluded to in the discussion on magnetism), for example, is an observable which affects physics greatly, but cannot be reduced to locomotion. The scholastic has no difficultly in incorporating this into the form of the particle. Is it an undue problem for the mechanist? Can he just expand his view of physics to 'matter in locomotion plus some other things'? The problem is that as soon as he starts adding to what affects physics, and acknowledges that this can differ from one fundamental particle to another, he needs to expand his description of matter, to include various other non-geometrical properties; and once he does so, he is in danger of

reintroducing the scholastic concept of form. He will call it by a different name, of course, but nonetheless mean the same thing by that name.

The fourth objection is that the mechanical physics, contrary to the picture of scholastic physics, made quick and rapid progress. However, once again, this is not a direct refutation of scholastic philosophy. Unsurprisingly, the best physicists latched onto Newton's, Boyle's, and Galileo's metaphysics. Since at first this seemed to be fully consistent with Christianity, particularly Protestant Christianity, there was no religious resistance against this movement. On the other hand, those scholars who followed the scholastic philosophy, like Aristotle and Aquinas, were not so interested in mathematics, and moved away from natural philosophy as a whole. There were three fundamental challenges facing the Medieval scholastic scientists: firstly, the mathematical methods required to properly analyse their system had not yet been developed; secondly the experimental methods that uncovered physics at a deep enough level for the scholastic principles to become apparent were not sufficiently advanced (indeed, classical mechanics was required as an intermediate step to construct and analyse the experiments that led to quantum mechanics); and thirdly the scholars had the wrong mindset: they were not thinking about these problems in the right way. The last of these was obviously the fault of the scholastic natural philosophers; the last two were not, and none of them stand against the intellectual system of scholastic metaphysics. Thus, possibly by historical accident, the vast majority of physicists became mechanists. Thus, with the most important key unveiled by Newton and some great minds perusing the problem, mechanical physics made rapid progress. However, this just shows that mechanism is consistent with classical physics. It does not show that scholasticism is inconsistent with modern physics, which is what is needed to draw the conclusion that scholasticism is false. That nobody succeeded in merging scholastic philosophy with modern physics does not prove that it cannot be done, particularly since nobody really tried.

The fifth objection was that causality is central to the metaphysics, but plays no real role in the mathematical approach of modern physics. However, as already shown, the principle of causality is encoded in physics, just in a more obscure way. Final causality, in the Aristotelian sense, is a necessary counter-part to efficient causality, and describes that only a limited number of effects can arise from an initial state. The early mechanists focused on design rather than final causality, but design is not a key component of scholastic metaphysics. Thus this objection also fails.

The sixth objection is that the search for final causes led people in the wrong direction. However, this does not mean that the idea is false; it may be that the search was done badly, or that this is not the right method to uncover science. But that does not mean that the concept of final causality is false; after all, it is only a small part of scholastic metaphysics, so even in scholastic terms an over-emphasis on teleology at the expense of investigating efficient, formal and material causes is likely to lead to problems.

The seventh objection was that physics should be performed without any metaphysical presuppositions. Were it possible, I would have some sympathy with this view. Although metaphysics is more fundamental than physics, in that fundamental physics should be derived from metaphysics, in order of our own discovery we come to physics first, and then need to draw out metaphysical conclusions by induction, falsification, or other means. There are, however, several issues with this. Firstly, the experimental method makes certain presuppositions about the regularity and nature of the universe. It is therefore, impossible to do science without making various metaphysical assumptions. The difficulty is that it is not just one set of coherent assumptions that are consistent with the scientific approach, but a large number of them. However, most scientists naturally are led into one set of working assumptions; most often those of mechanism because those were the assumptions that were held by the early developers of the scientific method, and were passed onto their disciples, and the disciples of those disciples, and so on down to our own teachers and mentors and then our generation; these assumptions are then absorbed uncritically into the scientist's

world-view. We are taught as fact in school that mechanism is both scientific and certain; how many people then will question this in sufficient detail? In the early days of modern science, the prevailing world-view was scholasticism, so naturally saying that we should not presume a metaphysical position was principally used as a weapon against scholastic metaphysics. However today, the prevalent world-views among scientists are mechanism and empiricism, so the warning against metaphysical presumptions should warn us against taking these too seriously. This objection is neutral in the question of whether scholastic, mechanical, or some other metaphysics should be adopted. It cannot therefore be used as a refutation of scholasticism.

The question then becomes can we do without metaphysics at all? Here the answer is no. Firstly, the various metaphysical propositions are mutually contradictory, so at most one of them can be right. On the other hand, the statement that there is no metaphysical foundation of physics, or that such a foundation exists but cannot be discovered, are both metaphysical statements, and thus self-refuting. Furthermore, metaphysics is important, because as well as physics, metaphysics provides the intellectual foundation for both religion (including atheism and agnosticism) and ethics, so this is not an issue which we can sensibly ignore. Finally, I note that many scientific advances have been inspired partly by metaphysical considerations (as well as partly from experiment), including Aristotle, the process leading up to Newton, and also Einstein's theories of relativity. The most notable exception to this was quantum mechanics, which emerged solely from experimental necessity, and caught the philosophers by complete surprise, though even here, as the field matured, underlying symmetry - i.e. philosophical - considerations became as important as experiment in shaping the development of the theory.

Finally, we have the objection that some aspects of classical physics contradict the traditional scholastic philosophy. The most notable example is the vision of space and time suggested by Galileo and Descartes (and their predecessors) and perfected by Einstein, Hilbert and Minkowski. Here the success of the geometrical view of space and time, its necessity in understanding gravity, and its inconsistency with the original scholastic view means that the scholastic understanding of space and time has to either be abandoned or modified. However, this is not a major part of scholastic metaphysics, and only a small modification is needed to bring the whole system up to date.

The second major issue which scholasticism faces when confronted with classical physics is that if we define a state in terms of both location and momentum (or velocity), then the principle of inertia implies that it is possible to effect a change (in position) without an external mover. One solution to this is to say that the identification of a state (as used in the principle of causality) only refers to the momentum. Different momenta states would then be the different potentia of a particle; and once we have the potentia we can start to distinguish form and matter. This allows us to maintain the principle of causality, and with it the rest of scholastic metaphysics, but at a considerable cost: we have removed location from the underlying metaphysical principles, which is problematic because both location and momentum need to be specified in classical physics. This problem can, however, I think be overcome by making additional minor modifications to scholasticism, although I do not intend to do so here because the problem is changed considerably once we introduce quantum mechanics so addressing it now in detail would be too much of a distraction.

An alternative is to note that scholasticism allows for the principle of natural motion, i.e. motion caused by the internal nature of something rather than something external; for example fire naturally travels upwards, or earth naturally travels downwards. As an explanation of gravity, this is utterly wrong; however it is quite close to the classical understanding of inertia and thus it is possible to adapt this idea to account for inertia by saying that material beings have a natural tendency to move at the same velocity from one potentia (specifying location) to another. Efficient causality links one particle state to another. So the efficient cause of a particle state described by a particular location and momentum

would be the same particle but in a different state (i.e. the same momentum but a slightly different location). The final cause of a particle state (in the absence of external forces) is a particle state with the same momentum but a change in its location (determined by the particle's velocity).

So to summarise, the early modern philosophers and scientists failed to present arguments against the premises or reasoning of scholastic metaphysics, so they failed to do the work which would have directly refuted the older philosophical system. The arguments they did present were partially based on misunderstandings, and otherwise for the most part completely failed to address the question: none of them present a refutation of the older metaphysics. At best, that mechanical arguments made more progress in physics than scholastic arguments did makes mechanism more likely; but we can only rule out scholasticism if all forms of it can be found to be inconsistent with the mathematical physics, and the early modern scientists and philosophers did not show that. We can, I think, state that the medieval philosophy needs some modification to be consistent with classical physics, but for the most part the modifications which I have identified are relatively minor. The only issue which represents a serious challenge I can think of is the dichotomy between momentum and location.

Although the mechanical world-view has been unanimously accepted by almost all of our intellectual elite, in almost every field, in terms of classical physics, its superiority over the one other consistent system we know of is marginal at best, and basically comes down to the question of whether scholasticism can be suitably modified to incorporate the mathematical elements of the new physics.

However, philosophically, the old system is considerably superior to the new one. There are numerous problems in enlightenment philosophy, which takes the mechanical world view as its base point, for example involving the philosophy of mind (the mind/body problem), ethics (the is/ought dilemma), and the problem of knowledge, which are simply not present in the older system (Feser, 2015). The Thomist version of scholastic metaphysics presents a coherent and complete world view, with the only unresolved issues matters of minor details, mainly caused by broader than necessary definitions which can be made more precise once the link to physics is re-established. So if it were not for the issues reconciling it to classical physics, scholasticism should be preferred.

Such was the position leading into the twentieth century; but with the empirical backing physics was obviously given priority over the philosophical issues, and one or other of the various forms of enlightenment philosophy generally accepted, with their negative implications for religious belief.

But then there was a game changer. Classical, mechanistic, physics was proved by experiment to be wrong. The most fundamental theory of physics is not mechanistic. That means that the philosophy built on the mechanistic picture, Descartes, Hume, Locke, Kant, Mill, Marx, Burke, Russell and all the rest who either wholly or partially assumed the mechanistic world view, are cast into considerable doubt and have to be discarded, at least until we have resolved the physics. Their premises were wrong, and therefore their conclusions have to be questioned. It changes everything. Once again, we have to tear up all the old work in philosophy, and start again. And this affects almost every idea and principle that our modern world is built on.

CHAPTER 9

Statistical Approaches to Physics

Before addressing quantum mechanics, I would like to mention another, important, area of physics: systems with large numbers of particles. The picture developed here will be largely unchanged whether we discuss classical or quantum mechanics; indeed, statistical mechanics, the branch of physics which deals with these systems (which underlies and justifies the older – but still important branch of physics known as thermodynamics) is something of a bridge between quantum and classical mechanics. I will, however, discuss this subject in my own way; to try to bring out the main points which I will need later. I will not discuss statistical physics in any detail not because it is not worthy of it (the field is of the utmost importance), but only because it would be too much of a digression even for me. My focus is in the area where physics and metaphysics intersect, which is fundamental physics; statistical physics builds on fundamental physics to calculate the properties of systems of large numbers of particles (for example gases, liquids and metals). It is a cornerstone of the modern field of condensed matter physics (the study of solids), and enables us to calculate the properties (such as its colour, electrical resistance, hardness, brittleness and so on) of a solid, for example a metal, plastic or glass, from knowledge of its underlying structure. I will need to refer to statistical physics in a few places, so it is worth providing a brief outline of the approach.

9.1. Statistical Physics

Suppose that we want to study a gas. One way of doing it is to carefully measure the position and momentum of every gas particle in the room, carefully thread that information into Newton's laws of motion, and solve the system. In practice, this is not a good idea. Two reasons why it is not a good idea are that it is far too much work since there are far too many particles for us to cope with, and secondly it gives us far more than we need in any real-life situation. When studying a gas, we are most interested in average quantities – pressure, temperature and so on, and going down to the molecular level distracts us from this.

The alternative is to take a statistical approach (Schroeder, 1999; Reif, 1965; Atkins, 2010; Stowe, 2014; Fermi, 1936). We assume that the location and momentum of each molecule of gas are statistically independent of each other and distributed according to some fixed function which is the same for every molecule in the room (fair enough if, for example, we are looking at say a gas of pure oxygen or a gas of pure nitrogen; the model can be refined if we have a mixture of different types of molecule, or there is some correlation between the different particles, though in practice the naive picture where we assume that they are all the same works very well). We can then treat the system probabilistically.

So the question is, what is the required probability distribution? Let's start with a system of two particles, particle A and particle B. We assume (again, for the first simple approximate model; obviously when physicists apply these methods to real-life systems they will be more sophisticated) that the particles only interact through collisions with each other, and, if they are in a box, with the walls of the box, though collisions with the box walls on average do not change the energy in the system, so particles are just as likely to gain energy from the outside world as they are to lose it. We also assume that the temperature of the box is kept constant. The position, momentum etc. of these particles for some observable

(let's say speed) are described by probability distributions P_A and P_B. We can write the total probability as $P_{AB} = P_A P_B$. There is another assumption here: that the energies of the two particles are independent of each other, which is a little dubious in this two particle system (we require that the energy of each particle is randomised enough through collision with the box walls that they lose any dependence on each other), but is more reasonable once we extend this picture to a huge number of particles. Now we should consider the sort of things that these probability distributions can depend on, and in classical mechanics there are only really three possibilities: the energy (E) of each individual particle, the momentum ($\vec{\mathbf{p}}$) of each individual particle, and the temperature (T). However, the only form of energy present is kinetic energy, and momentum is a vector quantity while probability is a scalar: we need some function of the momentum which gives a scalar, and the only option is the scalar product, $\vec{\mathbf{p}} \cdot \vec{\mathbf{p}}$. But this is just proportional to the Kinetic energy. So, the probability distribution is a function of just energy and temperature; $P_A(E_A, T)$ where E_A is the energy of a particle. This function gives us the probability that if we look into this system, knowing nothing else about it except the basic set-up, that we will discover particle A has an energy E_A.

Now, we know that energy is conserved in a collision, so $E_A + E_B$ is constant for these events. We also know that after a collision, energy is transferred from particle A to particle B. Let us say that the transfer of energy during a collision is ΔE, and that this is small compared to the total energy of the each particle. In Newtonian physics, the likelihood of a small energy transfer between two colliding particles doesn't depend on the amount of energy each of them have, so there is an equal chance of transferring a small amount of energy from particle A to particle B in a collision as there is from transferring from particle B to particle A. Obviously, eventually, we have to worry that the maximum energy we can transfer from particle A is E_A, but for very small transfers this is not a concern. So, given the initial state (and neglecting temperature dependence for now), the statement that particle A is just as probable to gain a small amount of energy in the collision as lose it can be written as,

$$P_A(E_A + \Delta E \setminus E_A, E_B) P_B(E_B - \Delta E \setminus E_A, E_B) = P_A(E_A - \Delta E \setminus E_A, E_B) P_B(E_B + \Delta E \setminus E_A, E_B). \quad (294.1)$$

$P_B(E_B - \Delta E \setminus E_A, E_B)$ means the probability that particle B will have energy $E_B - \Delta E$ after the collision conditional on the initial energies being E_A and E_B (and also conditional on the physical theory we use to describe the interaction being correct). But the behaviour of a given state in Newtonian mechanics (and, indeed, this survives into quantum physics) doesn't depend on past history: it doesn't matter how a particle got into a particular state, all that matters in the equations is what state it is currently in. In a system of many particles, a similar condition applies if the system is in what is known as equilibrium. Equilibrium means that the probability distributions governing the distribution of energy in the system are independent of what happened in the past. P_A will be the same regardless of the previous state before the collision. Thus we can say that

$$P_A(E_A + \Delta E \setminus E_A, E_B) = P_A(E_A + \Delta E \setminus E'_A, E'_B), \quad (294.2)$$

where E'_A and E'_B represent the energies of some other possible initial state that collided and led to the same final state. All we require is that the conservation of energy is satisfied, i.e. $E_A + E_B = E'_A + E'_B$. In other words, if the system is in equilibrium, we can drop the condition[1], and write that

$$P_A(E_A + \Delta E) P_B(E_B - \Delta E) = P_A(E_A - \Delta E) P_B(E_B + \Delta E). \quad (294.3)$$

We can now use a series expansion in ΔE to give,

[1]Note that the probability is still conditional in the sense that it depends on the various assumptions we built into this model; I have suppressed mention of these in my notation.

$$\left(P_A(E_A) + \Delta E \frac{dP(E_A)}{dE_A} + \ldots\right)\left(P_B(E_B) - \Delta E \frac{dP(E_B)}{dE_B} + \ldots\right) = \left(P_A(E_A) - \Delta E \frac{dP(E_A)}{dE_A} + \ldots\right)\left(P_B(E_B) + \Delta E \frac{dP(E_B)}{dE_B} + \ldots\right). \quad (295.1)$$

This needs a little clarification. $\frac{dP(E_A)}{dE_A}$ means the rate of change of the probability density P with respect to a small change in the energy – if we imagine the probability density as a function of energy, this gives the gradient of that function at a particular point. If we zoom in close enough to the function, we can neglect its curvature and it will seem to be a straight line. Using the equation of a straight line, we can say that the function at $P_A(E_A + \Delta E)$ will be close to the value at $P_A(E_A)$ plus a small correction – and that small correction will just be ΔE times the gradient of the line. Obviously, there will be corrections to this coming from the fact that the function is not a straight line these are contained in the $+\ldots$ in the equation – but these will be small, and the smaller we make ΔE in comparison to the energy of the particles the less important these corrections will be, so in practice we need not consider them. The above equation simplifies to for infinitesimal ΔE,

$$\Delta E\left(\frac{dP(E_A)}{dE_A}P(E_B) - \frac{dP(E_B)}{dE_B}P(E_A)\right) = 0, \quad (295.2)$$

which can be rearranged to give,

$$\frac{dP(E_A)}{dE_A}\frac{1}{P(E_A)} = \frac{dP(E_B)}{dE_B}\frac{1}{P(E_B)}. \quad (295.3)$$

The left hand side of this equation just depends on the energy of particle A, and the right hand side just on the energy of particle B. So the left hand side of this equation is some function of E_A, and the right hand side the same function of E_B; let's call this function $f(E_A)$ or $f(E_B)$. E_A and E_B are linked by the conservation of energy equation, so if the total energy of the whole system is E, then we have

$$\frac{dP(E_A)}{dE_A}\frac{1}{P(E_A)} \equiv f(E_A) == f(E_B) = f(E - E_A). \quad (295.4)$$

The only function that satisfies this condition for general E and E_A is a constant, which I will write as $\frac{-1}{k_B T}$. Thus, we obtain,

$$\frac{dP(E_A)}{dE_A}\frac{1}{P(E_A)} = \frac{-1}{k_B T}, \quad (295.5)$$

and the solution to this equation is

$$P(E_A, T) \propto e^{-\frac{1}{k_B T}E_A}. \quad (295.6)$$

Which means that the overall distribution is,

$$P(E, T) = P(E_A, T)P(E_B, T) \propto e^{-\frac{1}{k_B T}(E_A + E_B)} = e^{-\frac{E}{k_B T}}. \quad (295.7)$$

(295.6) gives the probability that the individual particle A has energy E_A. k_B is known as Boltzmann's constant, and the overall distribution as the Maxwell-Boltzmann distribution. The temperature T is introduced into this equation based largely on experimental observations: we know that at absolute zero temperature everything freezes in the lowest energy state, while if we gradually raise the temperature we start getting higher energy states, vibrations and so on. The form of the probability distribution means that at $T = 0$, the only energy state that can give a non-zero probability is at $E_A = 0$, which is what we see in experiment. Of course, we can ask why do we use a function E/T and not (say) E/T^2 which has the same effect, but this question slightly misses the point. We are free to define how we measure temperature and define the temperature scale; absolute zero is a physical point which our formalism needs to respect, and we need the physical quantity and our temperature scale to be related in a continuous and monotonic way to avoid confusion,

but otherwise we are free to define the quantity we call temperature however we see fit. So equation (295.6) can suffice as a definition of the measurement scale for temperature. We can then tune our thermometers so that the numbers on them correspond to the definition in this equation.

Although I have just provided a quick and dirty motivation of the relation, the distribution can be derived more rigorously, and is quite general. (295.7) can be generalised to a gas of any number of particles; I just used two because its the simplest case. However, since the expression is probabilistic, to convert the distribution into testable predictions the more particles we have in the system the better the results will be. Expected deviations from the calculated average in measured quantities scale as the inverse square root of the number of particles; to make a precise prediction we need this deviation to be as small as possible, either by having lots of particles, or averaging over a large number of independent experiments.

So what this equation tells us is that the probability of finding a particle with an energy E in a gas at equilibrium is proportional to the exponential of the energy. How does this help us? A great deal. For example, if we want the average kinetic energy of particles in the system, we write,

$$\langle E \rangle = \frac{\int dE E e^{-\frac{1}{k_B T} E}}{\int dE e^{-\frac{1}{k_B T} E}}. \tag{296.1}$$

The average velocity is

$$\langle \vec{\mathbf{v}} \rangle = \frac{\int dE \vec{\mathbf{v}} e^{-\frac{1}{k_B T} E}}{\int dE e^{-\frac{1}{k_B T} E}}, \tag{296.2}$$

and since kinetic energy is a known function of the velocity, a simple change of variables re-expresses this in a form we can easily solve. The pressure on the walls of the box can be computed in terms of the change of momentum of particles hitting the wall of the box, which can be computed from the average velocity and mass of the particles. The various thermodynamic quantities, such as responses of the pressure to changes in the temperature can also be calculated from this probability distribution. It is possible from this equation to recover everything that was known about gases, from the studies of Boyle onwards, and all the thermodynamics of gases: all the experimentally derived laws of gases that were painstakingly collected from the seventeenth to nineteenth centuries can be shown to follow just from four assumptions: that energy is conserved; that each particle of the same type should be treated in the same way; that the distribution for the energy of a particle only depends on the energy, temperature, momentum, etc. and other properties of that particle, is independent of the states of any other particles; and that in an equilibrium system the past history of the system does not affect the probability distribution for the current state[2].

These assumptions are quite general, and apply to every area of physics, not just the billiard ball model of gases. Even within the theory of gases, we can refine the model by adding the various gravitational and electromagnetic potentials into our expression for the energy, by allowing for different types of particle, and so on. But the basic formula $P(E) \propto e^{-\frac{1}{k_B T} E}$ remains, depending only on the assumptions that energy is conserved in particle interactions, that the system is in equilibrium, and that distributions for each individual particle are independent of each other. It applies in the physics of solids, in fluids: from the interactions of fundamental particles to the various metals and plastics we see around us: as long as there are enough particles and they are in equilibrium, the same basic equation applies. All we need is an expression for the energy of the system. The principle survives the transition to quantum physics, and our modern understanding

[2]I also supposed that the energy transfers were infinitesimal, but in a more rigorous derivation this assumption is not required.

of solids (which matches experiment to incredible accuracy) is based on the combination between this relation and quantum field theory, using the quantum field theory expression (based on operator arithmetic rather than simple functions) for the system's energy.

A new feature introduced by statistical mechanics is that we can think about physics probabilistically rather than definitely. To express the state of the system of many different particles, we do not need to state each particle's position and momentum individually, but we can express the state of the system as a probability distribution, with the probability characterising our ignorance of the individual microscopic details of each particle. We don't need a precise knowledge of the underlying system to perform useful and accurate physical calculations. This idea that physics can be treated probabilistically rather than deterministically will be very important as we go on.

The other important point is that we could uncover the form of this distribution with just a few basic, and almost self-evident, assumptions (with conservation of energy reducible to the translation of the time coordinate axis), and draw many grand conclusions over a wide range of objects from it which previously had to have been induced from experiment. This is important, because it provides us with a link between metaphysics and physics. Metaphysics tells us what these assumptions should be; physics then uses them to frame a theory, and, best of all, a theory which can be related to experiment.

It may be objected that the derivation of the Maxwell-Boltzmann distribution requires that the probability of particle A transferring a lump of energy to particle B is the same as the probability of particle B transferring that energy to particle A. We are working under the assumption that we do not know precisely what energy A and B have – and our uncertainty is classified by a particular probability distribution. We know which particles are in the system; we know the dynamics that control the evolution of the system, but we do not know precisely what energy and momenta each particle in the system has, with the uncertainty originating from a lack of knowledge of the various interactions transferring energy from one particle to another: what we know is described by some probability distribution. So the Maxwell-Boltzmann distribution requires the assumption that the distribution of energy of A and the distribution of energy of B are such that on average the A is just as likely to transfer a small amount of energy to B as it is to receive it from B. But how can we show that this condition is always satisfied? The answer is that it is not always satisfied. We say that physical systems where this condition holds are in statistical equilibrium, and any system where it does not hold are not in statistical equilibrium. Note that the definitions I have used for *equilibrium* and *statistical equilibrium* are different. They are two different concepts, albeit related as I will discuss below. Strictly Maxwell-Boltzmann statistical physics only applies to systems that are in statistical equilibrium. The study of non-equilibrium systems is a different (and much harder) branch of physics. Is this too much of a restriction to allow us to rely on it in practice? No, because there is another well-known physical law which states that over time, every physical system tends towards equilibrium. This means that if we close off a system, wait for a while, then it will be in statistical equilibrium and it will be well characterised by the Maxwell-Boltzmann distribution.

9.2. The Second Law of Thermodynamics

A many body system in classical mechanics is described by the position and momenta of all the particles at a given time. These numerous numbers can be grouped together into a single object, which I will label as $|\Psi(t)\rangle$. We may think of $|\Psi\rangle$ as a vector containing the momentum and positions of the particles as components, but this is in practice too simplistic (and impossible when we note that momentum and position have different units). Instead, each component of $|\Psi\rangle$ represents some suitable function of the momenta and position which provides a suitable representation of the coordinates for the problem we wish to solve: it is in principle possible to unpack the momenta and location of the particles from $|\Psi\rangle$ is some well defined way of our choosing. The only stipulation I will make restricting how this is done

is that we must choose these rules so that the overall normalisation of $|\Psi\rangle$ is unimportant: we can multiply $|\Psi\rangle$ by any factor and leave the extracted values of the momentum and position unchanged (this stipulation will ultimately represent some symmetry of the system; in practice it does not matter how we choose to parametrise our coordinates, so we are free to make this requirement or not as we choose, but there is an important point which is easier to explain if we require this). Changes in the state may be described by operators, something which takes one possible $|\Psi\rangle$ as input, and spews out another $|\Psi\rangle$ as an output. So we may write an operator relation as $|\Psi(t')\rangle = \hat{O}(t,t')\,|\Psi(t)\rangle$ (it is customary to write the equation in this way; however I do not want to restrict $\hat{O}$ only to linear operations; it would be stricter to write the equation as $|\Psi(t')\rangle = \hat{O}(t,t',\Psi(t))$). This equation is basically saying that we input the state $|\Psi\rangle$ and the relative times into the operator $\hat{O}$, and it will give us the new state $|\Psi'\rangle$. How this $|\Psi'\rangle$ is constituted depends on both the operator $\hat{O}$ and the initial state. I will restrict $\hat{O}$ by requiring it to be symmetric (if $|\Psi\rangle$ is real) or Hermitian (if $|\Psi\rangle$ is complex[3]), because it turns out that for physics every operator we need is either Hermitian or can be expressed as a function of a Hermitian operator.

Now the most important operator is what I will label as the pseudo-Hamiltonian operator, $\hat{H}$, which describes the time evolution of the system (*pseudo-* because I have simplified things a bit to have an easier presentation; this not the same as the Hamiltonian operator of classical mechanics; Hamiltonian because the Hamiltonian operator in quantum mechanics plays a similar role). This can, in principle, be constructed from the classical equations of motion for each of the individual particles. For an infinitesimal time δt we can write $|\Psi(t+\delta t)\rangle = |\Psi(t)\rangle - \delta t\hat{H}\,|\Psi(t)\rangle$. In Newtonian mechanics, the forces on each individual particle at a given time depend on the position and momenta at that time of every particle in the system; the new position and momentum of the particle a moment later can be derived from the forces and the old position and momentum (of course, this picture has to be modified to account for relativity theory, but such modifications can be done; for simplicity I will stick to Newtonian mechanics to describe the general ideas). Note that the precise form of the Hamiltonian will depend on how we choose to represent the coordinates within $|\Psi\rangle$.

An important class of operators are linear operators, represented by matrices (and given that $|\Psi(t)\rangle$ and $|\Psi(t')\rangle$ are of the same length square matrices). Here the new $|\Psi(t')\rangle$ is just the sum of the components of the old $|\Psi(t)\rangle$, so, writing each individual position or momenta as $|\Psi_1\rangle, |\Psi_2\rangle, |\Psi_3\rangle, \ldots$ we have $|\Psi(t')_n\rangle = \sum_m \wr_{nm}\,|\Psi_m\rangle$ for some coefficients $\wr_{nm}$ which depend on the operator $\hat{O}$. Now a feature of linear operators is that there are certain vectors, known as eigenvectors, which are unchanged by the application of the operator. $\hat{O}\,|\psi_i\rangle = \lambda_i\,|\psi_i\rangle$, where λ_i is known as the eigenvalue and $|\psi_i\rangle$ the eigenvector. I will sort the eigenvalues so that $\lambda_0 < \lambda_1 < \lambda_2 < \ldots$. One of the properties of Hermitian or symmetric operators is that the eigenvectors form a complete basis; that is to say that any vector can be expressed as a linear combination of the eigenvectors of the system. Thus we may write $|\Psi(t=0)\rangle = \sum_n c_n\,|\psi_n\rangle$. Now if we apply the pseudo-Hamiltonian operator a large number of times, we would find that (if the Hamiltonian were linear) $|\Psi(t=N\delta t)\rangle = \sum_n c_n(1-E_n\delta t)^N\,|\psi_n\rangle$, where $E_n\delta t$ represent the eigenvalues of $\delta t\hat{H}$. If we further suppose that the eigenvalues of the pseudo-Hamiltonian are greater or equal to zero (and less than $2/\delta t$), with $n=0$ corresponding to the $E=0$ eigenvalue, then we would find that the system, as N increases, would eventually find itself falling into a state

[3]A complex number is one which contains an part proportional to the square root of -1 – this part is called the imaginary part of the number; a real number is the sort of number we are used to which doesn't. Complex conjugation is the procedure where we flip the sign of the imaginary part, so $1+i$ becomes $1-i$ after complex conjugation. A symmetric matrix is one that is equal to its transpose, so you can swap the rows with the columns and you get the same matrix. A Hermitian matrix is the analogue of this for a matrix with complex coefficients, where you swap the rows and columns and take the complex conjugate of each element.

proportional to $|\psi_0\rangle$: all the other eigenvectors are suppressed. Once it was in this state, it would remain there and never shift from it (this is true for any state that is an eigenvector of the pseudo-Hamiltonian). In other words, the many body system would be directed towards a state which represents the lowest eigenvector of the pseudo-Hamiltonian. And this must be the case for any system with a linear pseudo-Hamiltonian: eventually, after enough time, it will collapse into an unchanging state given by one of the eigenvectors of the pseudo-Hamiltonian.

What if the pseudo-Hamiltonian has several eigenvectors with a zero eigenvalue? In this case rather than the state tending towards a fixed point, characterised by a single eigenvector, it may tend towards a fixed orbit, where it moves from one to another (particularly if it is perturbed from time to time). But it will still be stuck on a limited number of states.

Now, of course in real life, the pseudo-Hamiltonian operator is usually not linear (or, expressed better, we cannot choose a representation of the coordinates so that the corresponding pseudo-Hamiltonian is linear). The discussion in terms of eigenvectors and eigenvalues is only valid for a linear operator. This actually changes less than you might think. It still may be the case that there still exist some fixed points of the system, where $\hat{H}|\Psi^*\rangle = \lambda|\Psi^*\rangle$ for some scale factor λ. In this case, if the system is at any time in the state $|\Psi^*\rangle$ it will remain in that state forever. These fixed points may be split into three classes: relevant, irrelevant, and mixed. A relevant fixed point is one where if the state is a small distance away from $|\Psi^*\rangle$ it will evolve towards $|\Psi^*\rangle$[4]. An irrelevant fixed point is one where if it is a small distance away from $|\Psi^*\rangle$, it will move away from $|\Psi^*\rangle$, and a mixed fixed point is one where if it is displaced in some directions it is relevant and in other directions it is irrelevant. Now the system must either move towards $|\Psi^*\rangle$, move away, or stay the same distance from it. If it stays the same distance, then either it must remain in the same place, i.e. another fixed point, or it orbits $|\Psi^*\rangle$ at some fixed radius: this I will label as a fixed orbit, and will discuss again later. Whether an operator has fixed points depends on the operator: some do, and some don't.

Suppose that we have a pseudo-Hamiltonian with one relevant fixed point and no fixed orbits, for a compact system (i.e. with a finite number or countable infinite number of possible states: for example, if the positions of the particles are constrained in a box, so there is a limited region in which they exist). We start the system anywhere. It will evolve, moving around the phase space (i.e., the possible states it can have). If the trajectory describing the motion of $|\Psi\rangle$ does not repeat itself (i.e. a fixed orbit, which we have forbidden), then eventually it will come into the vicinity of the relevant fixed point, and then it will be trapped: it will be sucked into the fixed point and unable to escape. In other words, the system has a tendency towards being at rest in one particular state. Suppose there are many fixed points: eventually it will fall into one of them; which one depends on where the system starts. Suppose there is a fixed orbit; eventually the system will fall into it: rather than be stuck in a single state, it will become stuck in one of a small number of states, but the final result is the same: the system is trapped. It cannot go everywhere in phase space; it is restricted to a small region.

What if there are no relevant fixed points or orbits in our pseudo-Hamiltonian? Have we escaped the trap? The system is now what is known as chaotic, which means that a tiny difference in initial state will eventually lead to a large difference in the state at the time of measurement. However, even in chaotic systems, there are things known as attractors. These can be thought of as fuzzy orbits. The state of the system is not trapped on the orbit, but it is trapped in the vicinity of the orbit. Once again, the system will be eventually remain in just a small region of phase space. Pseudo-Hamiltonians which are

[4]To discuss 'evolve towards' we need some concept of 'distance from'. There are many ways to define this; perhaps the simplest is to say that the distance between a state $|\Psi\rangle$ with individual coordinates Ψ_n and $|\Psi^*\rangle$ is $(\langle\Psi| - \langle\Psi^*|)(|\Psi\rangle - |\Psi^*\rangle) \equiv \sum_n |\Psi_n - \Psi_n^*|^2$. This is greater than zero. Evolve towards means that it decreases in time; move away means that it increases.

almost linear (that is that the non-linear terms are small in comparison to the linear terms) will often lead to attractor states, with the centre of the attractor being the orbit formed from the eigenvectors of the pseudo-Hamiltonian with the identical smallest eigenvalues. We split the pseudo-Hamiltonian into a linear part and the non-linear part. The linear part strongly forces the system to head towards the fixed orbit. The non-linear part is weaker, but doesn't care what happens to the system. If the state is on the orbit, then the linear part of the pseudo-Hamiltonian is content and does nothing, but the non-linear part will move it around by a small amount (since the non-linear contribution to the pseudo-Hamiltonian is small). If it strays too far from the orbit, then the linear part of the pseudo-Hamiltonian will get upset and push it back towards the orbit, while the non-linear part will be too weak to object. Thus the state will be trapped in the vicinity of the orbit, and we have an attractor. As I mentioned, the form of the Hamiltonian depends on the choice of the coordinate system, and all describe the same physics. If the pseudo-Hamiltonian is 'almost linear' for *any* of these possible parametrisations (subject to the single restriction I described earlier), then the system will be described by an attractor.[5] This seems to be the case for most physical systems.

So, these are the ways in which a many body system may evolve after a long period of time, depending on the pseudo-Hamiltonian: it may be trapped in a fixed point, trapped in a fixed orbit, trapped in an attractor, or (and only if the Hamiltonian is one of the rare, at least for physicists, truly chaotic Hamiltonians, with no attractors, fixed points or fixed orbits) free to go to any state. Note the predominance of the word 'trapped'. In time, many body physics evolves towards certain states, known as fixed states (fixed may refer to an orbit or the vicinity of an attractor). If it is a long distance from the attractor, then it will move towards it. If it is in the vicinity of the attractor, it will remain there. This may be expressed in various ways. For example, we can define the quantity 'entropy' as a measurement of how far the state is from the attractor[6]. The closer the state is to the attractor, the larger the entropy. We can then summarise this discussion by saying that 'entropy never decreases with time (at least for almost linear pseudo-Hamiltonians), and for all but the most carefully constructed processes will increase.'

There is obviously a close relationship between this discussion and the well known second law of thermodynamics. The second law, which has numerous equivalent definitions, is not usually expressed in this way[7]. Nonetheless, if you will forgive me for the misuse of a well known principle, I will call this the second law of thermodynamics: a many body system will always, after enough time, reach an fixed state. This has three important consequences which I want to highlight. Firstly, time has a direction, even if the pseudo-Hamiltonian itself is symmetric with respect to time (i.e. makes no distinction between going forwards or backwards in time). Secondly, if a physical system is in a fixed state, then it may have been going on indefinitely or only recently set into motion (i.e. by some factor external to the system) – we cannot tell. However, if it not in a fixed state, then it must have been set in motion in the relatively recent past; otherwise it would have reached the fixed state. It cannot have been in existence for an infinite time in the past. Thirdly, there is an in built tendency in most natural systems (as long as the external conditions remain constant)

[5]There may also be pseudo-attractors, where the system remains in the vicinity of the orbit for much time, but it may escape along certain paths; I lump these in with the 'truly chaotic' states with no attractor, fixed point or fixed orbit.

[6]Note that this definition of the word entropy is non-standard. It is usually defined in terms of the number of equivalent ways a state can be ordered, the highest entropy state the one with the most equivalent expressions one can make for $|\Psi_*\rangle$, i.e. if you were to through the universe in a state at random it would be the most probable state.

[7]There are about four or five equivalent definitions which are taught to physics undergraduates and usually immediately forgotten; the one I recall best is that 'No engine is more efficient than one which consists of two reversible adiabatic steps and two reversible isothermal steps;' and the others are of similar memorability. The unofficial, but less rigorous, definition, 'Entropy never decreases,' is easier on the tongue.

towards a particular, fixed, state. This is sufficiently obvious from observation that it was noted by the ancient natural philosophers. Although they would not have described it as I have done, and obviously lacked the mathematical machinery to do so, they referred to what I have labelled as the fixed state of the system as the 'final cause' of the system (or at least this is one of the several meanings which were represented by the phrase 'final cause'). Once we unpack the two different languages, the concept is precisely the same: that many closed complex systems naturally tend towards an end which is stable in time, and this end is the final cause. Therefore the possible existence of final causes (in this sense of the term) for at least some physical systems is built into mathematical law. No objection to this sense of the final cause raised by an enlightenment philosopher is valid.

All this, of course, depends on the pseudo-Hamiltonian being 'almost linear.' There are, however, other arguments for the convergence for physical systems.

Suppose, for example, we consider a system of a large number of particles; too many for us to measure each one's properties individually and precisely. We will therefore have uncertainty over the state of the system, and that uncertainty means that the best we can do is express the system using probabilities.

By saying that a particle is in a given state (in classical physics), I mean that we precisely specify the location, momentum, energy, and anything else of importance to the physics. Each specific set of values for one of these parameters corresponds to one possible state, which I will denote as s. Change the momentum (for example), and the particle will be in a different state s'. The state of the system, which I will denote as S, is simply the amalgamation of each individual particle state. So S means that we have one particle with this momenta and location, another particle with that momenta and location, and so on. The particles themselves are treated as indistinguishable, so it doesn't matter which particle has that momenta and location; if any one of them does, then we are in the specified state.

We do not know what the state of the system is at any given moment of time. We have uncertainty, and that means that there is probability. We can express the probability that at a time t the system is in a state S as $\Psi(S,t)$. Now we can think about how the system evolves in time. In a deterministic system (such as classical mechanics), we have an operator $\hat{\mathcal{O}}_{S',S,t'-t}$ which tells us that if the system is in a state S at time t, then in the next instant (at time t') it will be in a state S'. In classical mechanics, this operator describing the time evolution of the location and momentum of the particles is a representation of Hamilton's equations; in non-relativistic quantum mechanics, it will be a representation of the Schrödinger equation (albeit acting on a probability amplitude rather than directly on the probability); there are also similar operators in quantum field theory and even a fully worked out version of Bradwardine's mathematical expression of Aristotelian physics. We can apply this procedure to every state, which means that we can also apply it to the probability: the probability $\Psi(S',t')$ is the sum of all the probabilities $\Psi(S,t)$ such that the state S evolves into the state S'. Note that although I have been discussing deterministic systems, we can easily extend this argument to indeterminate physical theories by expressing $\hat{\mathcal{O}}_{S',S,t'-t}$ as representing the probability that the system will evolve from state S to state S'; the argument becomes a little more involved for an indeterminate physics, but the same basic principles apply.

Every physical observable for a particle can be computed from knowledge of the state s, since all the physically relevant information about the particle is described by our representation of the state. So, for example, there is a particular value of the energy E_s associated with that state. Our uncertainty of what the state s is is characterised by the individual particle probability distribution $\psi(s,t)$, and the system distribution (if these probabilities are independent of each other; this assumption is not necessary for the argument, but I will make it to keep things simpler) will be given the probabilities for each particle multiplied together. While we cannot specify what the energy of the particle is, we can compute the expected value (which corresponds to the average value if we repeat this process for a large

number of particles whose distribution between the states is proportional to $\psi(s,t)$): for each possible state, we multiply the energy of that state by the probability of being in that state, and sum over the states. In symbols, $\langle E \rangle = \sum_s E_s \psi(s,t)$.

If we know the time evolution operator $\hat{O}$ and the current probability distribution, can we compute the distribution at the next instant in time? This is given by the formula $\Psi(S',t') = \sum_S \hat{O}_{S',S,t'-t}\Psi(S,t)$. The operator $\hat{O}$ represents the physics of the system; for example it might encode Newton's laws of motion. Now every expression of physics we know of satisfies two conditions: firstly, that it only depends on the state of the system at that particular moment in time, so it does not depend directly on the past history of the system (the principle of locality); secondly that every possible state for a particle is in principle accessible. Apply exactly the right forces in the right order and for long enough time, and you can move a particle from any given state to any other possible state. Mathematically, these are the defining properties of what is known as a Markov chain, a phenomena which was well-studied during the twentieth century. I have been discussing two times, t and t', but one can, of course, repeat this procedure as often as we like, and we get a chain of probability distributions, one for each instant in time extending from the far past to the far future. One result of these studies into Markov chains is that they always converge to a fixed distribution, $\Psi_*(S)$ (Kennedy, 2006). What this distribution is only depends on the details of the operator $\hat{O}$, and not on the initial state of the system. This means that, after enough time, the probability that the system will be in a state S will be given by this distribution $\Psi_*(S)$ (where $\hat{O}$ acting on the distribution Ψ_* will give the same distribution Ψ_*), as long as we don't interfere from outside and just let it evolve naturally. Not knowing anything else about the system other than the nature of the particles within it and their interactions, the probability of there being a gas molecule in a particular small region of space in a particular small region of time remains the same no matter when we look at the system – as long as we leave long enough between 'looks' so our knowledge from the previous 'look' becomes diffused by the subsequent unknown interactions.

Correspondingly, the probability that each individual particle will be in a given state will also be fixed; if they are indistinguishable and their distributions are independent of each other they will be governed by the fixed distribution $\psi_*(s)$; if these conditions are relaxed then there will still be some distribution, only it might not be the same for each particle, and it might be harder to go from the system probability to the individual particle probability, but it still could be done. If there is a large enough number of particles, then the frequency distribution of the number of particles in each state will be proportional to the probability distribution. In other words (up to a minor variation which is less significant the more particles we consider), once the system has evolved for long enough to reach this fixed state, the number of particles in any particular state will not change over time. They will be different particles in those states at different times, but the number of them will be unchanged. This means that the macroscopic or average properties of the system, which depend on how many particles are in each state, will not (noticeably) change over time.

Since the particle reaches the fixed distribution no matter which state it began in, this fixed distribution is not dependent on the past history of the system: it is an equilibrium state.

Regardless of whether the underlying physics is determinate or indeterminate, as long as it satisfies the two conditions of locality and ergodicity, and there are enough particles that on we can neglect the minor statistical fluctuations (the fluctuations of an average quantity generally scale as the inverse square root of the number of particles), the frequency distribution of the particle states thus tends inexorably towards some fixed distribution. With these conditions (and if locality was violated in our universe its consequences would have been measured) the second law of thermodynamics cannot be avoided.

We can also say that this fixed distribution Ψ_* is also a statistical equilibrium distribution, in the sense defined above. Suppose that it is not a statistical equilibrium distribution.

This means (for example) that particle A is more likely to transfer energy to particle B than particle B is to particle A. This means that the energy of particle B will, on average over time, increase and the energy of particle A will decrease. This also means that the mean of the probability distribution of the energy of particle B will be increasing and the mean energy of the distribution for particle A will be decreasing; which means that the two probability distributions will change over time. However, the fixed distribution Ψ_* contains the distribution of the energies of the two particles, and this probability distribution does not change as the system evolves. There is a contradiction here, which means that our premise is false, which in turn means that the thermodynamic fixed distribution is one of the possible states of the system which are in statistical equilibrium. Note that equilibrium does not mean that the particles themselves are static, but only the probability distributions which describe our knowledge of the particles if we are unaware of the microscopic details concerning their detailed interactions. Each individual particle itself will be gaining and losing energy just on average after a large number of interactions it will gain as much energy as it loses.

9.3. Natural Selection

A good example of this is from biology, provided by natural selection. Each biological individual is defined by a particular genetic sequence. It is often said that two individuals may be viewed as part of the same species if they are capable of interbreeding, or, more precisely, their genetics are compatible with a given sample male or female (that a species is defined by its genetics rather than the physical attributes it possesses is obvious: a sterilised dog, even one which was neutered at birth or in the womb, is still a dog) I will give a different definition below, but this is good enough to start with. Every possible genetic combination represents a potential organism, and each of these have a certain probability of producing viable offspring, partly built on the genetic code and partly on external factors such as the surrounding environment. In practice, the overwhelming majority of these potential organisms are non-viable: they have no chance of breeding. We may thus have a probability distribution, $P(xy, E)$ which covers the space of all possible genetic codes, xy, and all possible environments E. Those individuals which have a large probability of survival in their particular environment will pass on their genes to the next generation and the genetic codes will survive a long period of time (as long as the environmental factors remain constant); those individuals which do not will not survive and their genetic codes will die out.

This may be modelled by a Hamiltonian system. The input state $|\psi(t)\rangle$ describes the probability that the various possible genetic codes are realized in practice in that particular generation. An element of the vector, say $\psi(t, 137)$ represents the probability that there is an individual with a genetic code corresponding to the 137th item on our list of all possible DNA sequences[8]. The output vector, $|\psi(t+1)\rangle$ describes the similar probability distribution for the next generation[9]. The pseudo-Hamiltonian, $\hat{H}$ describes the mechanics of the system: it includes the probability that each particular individual will breed given the environment and the other individuals present. Thus we will have the equation $|\psi(t+1)\rangle = \hat{H}\,|\psi(t)\rangle$, and if the external environmental factors, and therefore $\hat{H}$, are not changed significantly, this will gradually move towards one of a certain number of fixed points where $\hat{H}\,|\psi^*\rangle = |\psi^*\rangle$,

[8]For simplicity and ease of explanation, I am excluding the possibility that there are two individuals with a given genetic code; i.e. identical twins or clones, although the model can be easily extended to cover this.

[9]Of course, in real life, not every species has the same life span and therefore the same period between generations; again the simple model I am describing here would if used in practice be extended to make it more realistic – one can regard t as indicating time periods rather than generations, and it makes no real difference to the conclusions.

if such fixed points exist[10]. In each of them the majority of individuals will be clustered around the various maxima of the probability distribution P. If these clusters are widely enough separated (with an infinitesimal probability of occupation between them) we can distinguish the members of each cluster as the members of a particular species.

Now most of the time, the genetic code of the next generation strongly depends on the genetic code of the previous generation. For example (switching off mutations for the moment, and making the vastly oversimplification that each individual carries only one bit of information), if we have one man present with DNA strands X_1Y_1 and one woman with DNA strands X_2X_3 and suppose that they are capable of interbreeding. We describe this state as $|\psi(0)\rangle = |X_1Y_1\rangle + |X_2X_3\rangle$, meaning that the probability of there being an individual $|X_1Y_1\rangle$ in the initial generation is one, and $|X_2X_3\rangle$ is also one. The next generation will consist of individuals with genetic codes X_1X_2, X_1X_3, Y_1X_2 and Y_1X_3, and these will all be present with roughly equal probability. If they have two distinct children, then we may find $H|\psi(0)\rangle = |\psi(1)\rangle = \frac{1}{2}|X_1X_2\rangle + \frac{1}{2}|X_1X_3\rangle + \frac{1}{2}|Y_1X_2\rangle + \frac{1}{2}|Y_1X_3\rangle$. If, however, Mr Y_1X_2 is much better adapted to breeding than Mr Y_1X_3, then the next generation will be dominated by his children rather than Y_1X_3's What we cannot have is an individual $Q_{1000}T_{75}$, whose code is not contained in the parents.

Adding the possibility of mutations (where the genetic code is not exactly copied from one generation to the next, but random differences are introduced), or variable numbers of children in each generation, does not affect the overall picture: all this information is included within $\hat{H}$. As long as the mutations are not too severe, they will not severely affect the fixed points of the system: the individuals will still be clustered around the maximum values of the probability distribution P, the mutations will merely affect the distribution of individuals around that maximum, and in general a large rate of mutations will allow individuals to be further from the maximum than a smaller rate[11]. However, as long as the mutation rate is small enough, the clusters will still be separated from each other and we may still identify species. The only significant difference that mutations make is that they, in principle, allow the entire space of viable organisms to be sampled no matter the starting distribution (otherwise, we would be restricted to only those genes that we present in the first parents).

So given any initial distribution $|\psi\rangle$, and limited resources in the system so only a certain number of individuals can exist in each generation, we will find that those parents which are closer to the maxima of P will produce more offspring than those which are far from the maxima, so the next generation is better adapted to breeding than the previous one, until we reach a population that is well adapted. Mutations will attempt to move the population away from the maxima; natural selection will kill off the lineage of any organisms whose genes stray too far. The final distribution is entirely determined by the fixed point structure of $\hat{H}$ and the initial distribution (which only determines which of the fixed points of $\hat{H}$ is reached at the end), and it depends solely on the environment E. For a large enough sample of individuals, the final population distribution is not in any way random, despite the inclusion of random mutations into the theory. Given $\hat{H}$ and enough intelligence (neither of which we have in practice), we could predict what it would be as well as we can predict the distribution of gas molecules in a room. Note that, in an unchanging environment, a fixed point of $\hat{H}$ does not mean one species alone, but will contain a large number of individuals of various species, but with the number of individuals of each species roughly constant on average across large time scales (though there will be variations on smaller time

[10]If there are no such fixed points, then attractors or fixed orbits will give the same conclusions; only a truly chaotic system would differ, but observation suggests that the system is not chaotic as species are stable over long periods of time if there are no environmental changes.

[11]Note that there are two probability distributions at work here, $P(xy, E)$ the probability that an individual will breed which is contained within $\hat{H}$ and the population distribution $\psi(t, xy)$, the probability that the individual xy exists within the generation t.

scales). It means a biosphere where nature, with a variety of species, is in balance. However, which of the almost infinite number of species that could exist in a given environment are in practice populated depend on which fixed point of $\hat{H}$ we are in. Note that because $|\psi\rangle$ refers to a probability, this does not mean that actual populations will be constant from one generation to another; only that the average over a large number of generations will not vary significantly.

Natural selection is thus more 'elimination of the weakest' than 'survival of the fittest', and in particular it eliminates those possible genetic codes which lie between one species and another. How then do new species arise? There are two possibilities that I am aware of. Firstly, $\hat{H}$ may not involve completely stable attractors, i.e. it is possible that a trajectory may jump from one fixed point of $\hat{H}$ to another; if the area in the genetic states between species is only represented by vanishingly small probability rather than zero probability, then there is still a chance that a population might cross the barrier between one species and another; given enough time this chance would be realized. This is possible in chaotic systems. I do not know whether it is possible in this particular system. Secondly, and more importantly in my view, the environment is not constant. The particular fixed points of $\hat{H}$ depend on the environment. If there were a sudden shift of the environment (whether natural or man-made), then the fixed points of the system would change. Suddenly the population would not be in equilibrium; and it will rapidly thermalise to a new fixed point. Then, if the environment is suddenly shifted back, the population distribution will again be in turmoil, and will thermalise to a fixed point. If the shift in the external environment is sufficiently large, then it need not return to the same fixed point that it started in. This means that new species may arise.

An often cited example is antibiotic resistance in bacteria. The original distribution of bacteria, in the absence of antibiotics, covers a large range of genetic codes clustered around the maximum of the probability distribution. Some individuals will be resistant to the antibiotics (or such individuals will be only a small mutation away from the original distribution of the species), but as this offers no advantage in the absence of antibiotics, these will be few in number. We now introduce the antibiotics. The probability distribution changes; and the maximum is now centred around the antibiotic resistant strains of bacteria. The population distribution will suddenly be out of equilibrium, but it will quickly thermalise to this new maximum. Most of the bacteria will be killed; a few will survive and then populate the new distribution. Suddenly, all the surviving bacteria will be antibiotic resistant, and, as they multiply, the population will stabilise around a new fixed point which is resistant to the antibiotics. Of course, if we remove the antibiotics, given that there is no shift in species, the system will again be out of equilibrium and the population of bacteria will gradually revert back to its original distribution. This will, however, take a far longer time, as the population needs to be redistributed by the far slower process of mutation rather than the rapid process of elimination.

Is evolution by natural selection enough to create the diversity of life we see around us? Certainly most contemporary biologists believe so, and it is plausible although to my knowledge (which is admittedly not very extensive in this area) it has not been rigorously proved. It depends chiefly on whether all species present today and in the past are connected by a path through the landscape of genetic codes representing possible viable life forms linking us to that first DNA sequence capable of reproduction. By viable, I mean that they are in principle capable of living and reproducing, even if they are very poorly adapted to their environment. If we were to throw together a random genetic code, then in almost all cases the creature would be unable to live or develop for even a moment (for example, it might describe a mammal-like creature without a heart or which is unable to digest food). Note that this path is in the landscape of genetics, not physical attributes. It may be that a long necked giraffe is just a small mutation from a short-necked giraffe; if so, even if the physical intermediate state (a medium-necked giraffe) is non-viable, then it can

still be reached genetically. This is a difficulty I have with the usual presentations by the intelligent design movement and their concept of 'irreducible complexity': the question is not whether we can get from a bacteria without a flagellum to a bacteria with a flagellum in small physical modifications of the bacteria each producing some useful improvement, but whether the genetic code representing the flagellum bacteria is sufficiently close to that of a sustainable bacteria without the flagellum that it cannot reasonably be reached by various small mutations without passing through a non-viable or disadvantaged state. To prove 'irreducible complexity', one would have to demonstrate that the genetic code of the species is surrounded by a wasteland representing non-viable organisms and which is too large to jump in a single mutation. This may be the case, but it would have to be convincingly proven, which to my knowledge it has not been. I have, however, seen attempts to show that it could arise through a sequence of typical mutations (Pallen and Matzke, 2006)[12]. My knowledge of biology is far too scant to make a judgement either way in cases such as this (and even if I did, it would be worthless); if I were a betting man, though, my money would be on the supporters of evolution rather than those promoting irreducible complexity.

There are a number of important points to be made:

(1) Natural selection is not random. I have seen several theologians claim[13] that natural selection is incompatible with theism because of the random element to it. The argument is straightforward: God intended to create Mankind as a goal of the act of creation; evolution by natural selection is a random process and thus unpredictable; unpredictable outcome is inconsistent with prior intention; therefore God did not use natural selection in His creation of life. Of the three premises, I agree with the first, and disagree with the third, as it makes incorrect assumptions about God's means of foreknowledge (indeed the very word 'foreknowledge' is suspect, as we shall see later), and that it is unpredictable for us does not mean that it is unknowable for God. But it is the second premise which I wish to address here. Although it contains a 'random' element in terms of mutations (although each individual mutation arises from definite causes; it is only unpredictable to us because we are unaware of those causes), for a large enough population random does not mean wholly unpredictable, at least at the level of the population. The outcome of the process of evolution can be precisely determined by the fixed point structure of the Hamiltonian and the initial state. The fixed point structure of the Hamiltonian can be determined by the environment and the physical attributes of each species that result from the genomes. None of these are unknown to God, or outside His power. If a rational animal is possible (which is certainly true) and is the creature least affected by the environment and therefore most suited to survive (which can be disputed but is plausible) then, given enough time, one form of rational animal is pretty much certain to emerge. Indeed, as the processes behind natural selection, even the random elements, ultimately have their origin in physics, and since in theism physics is simply a mathematical description of how God sustains the world (as we shall see later), evolution by natural selection can both be random (in the sense of unpredictable) and ordained by God, as long as we understand what appears to us to be quantum mechanical randomness as being due to the freely chosen decisions of God. But I will discuss this in considerably more detail later.

(2) It is often argued by biologists that natural selection undermines theism because it provides an explanation of the diversity of living organisms which does not require divine intervention. This is also incorrect, because it presupposes a mechanistic

[12]See *http://www.evolutionnews.org/2011/03/michael_behe_ hasnt_been_refute044801.html* for one Intelligent design response to this work; there are, of course, also responses to this response

[13]For example see (Grudem, 1994), which, baring the odd scientific mistake such as this ad a few minor historical errors, is otherwise an excellent book.

metaphysics, i.e. that the laws of physics, which underlie natural selection, are independent of God. This is not a premise that a classical theist will agree with (indeed, I believe that mechanism, in its most straight-forward forms, is as incompatible with theism as it is with quantum physics). Just as the laws of physics describe the work of God, so does natural selection. Natural selection, therefore, in no way undermines the need for God once we accept a metaphysics which is consistent with theism (which we ought to do if we wish to discuss their consistency, otherwise we will committing the fallacy of assuming the conclusion). Whether by direct intervention, or through His general providence as described by the laws of natural selection; in both ways God is ultimately responsible for the diversity of life. The proposal that evolution by natural selection is inconsistent with theism is thus based on (incorrect) philosophical (and ultimately metaphysical) premises, and not on the underlying science. The argument used by the biological inspired atheists against theism is built on two premises: firstly evolution by natural selection, but secondly and crucially an underlying mechanistic metaphysics. But it is mechanism (the assumption that the laws of physics operate independently of God), not natural selection, which causes the difficulties with the idea of a theistic God. If, however, the theist is correct in saying that physics describes the acts of God in sustaining the universe, then ultimately evolution by natural selection is a description of how God's act of sustaining of the universe has directly led to the origin of the different species we see around us. There is therefore not even the slightest tension between the theory of evolution by natural selection and theism, although there is tension between evolution and those forms of deism where God intended to create mankind.

(3) This will make biologists squirm: natural selection is unavoidably teleological. I have described it in terms of equilibrium states of a pseudo-Hamiltonian (or, at the very least, meta-stable states); and thermalisation towards an equilibrium state is equivalent to proceeding towards a final cause (in the sense of 'final cause' used above). We have two choices: to abandon the unnecessary premise of the fathers of modern science that teleological explanations have no role to play in science; or to accept the most natural implications of the results of natural science. The second of these is by far the most preferable.

(4) It is sometimes argued that natural selection makes discussion of species difficult. If men emerged from an ancestor similar to the modern chimpanzee, at what point did it stop being a chimpanzee and start being a man, since the process of evolution is continuous? Where do we draw the line between chimpanzee and human? (You can replace chimpanzee with another creature if you believe that men emerged from something else - such as a common ancestor of the two modern species; it does not affect the argument). The modern boundary lines between species become blurred, and are only a temporal accident. This, however, is flawed. Given a particular environment, the species are defined as contingent regions of non-infinitesimal probability surrounding the maximum probability of the fixed point (or meta-stable states) of the pseudo-Hamiltonian. They depend on the environment (although should be stable under small enough changes of the external environment) and the particular fixed point of the Hamiltonian selected by nature. For each given environment, the species are well defined. Species are defined in terms of the equilibrium distribution (or the meta-stable states if there is no equilibrium). Intermediate creatures, the 'missing links', occur when the system is thermalising. They are neither of one species or another; as the definition of species depends on the equilibrium distribution.

(5) The individual with the maximum probability for passing on the genes within a given species is similar in several respects to a Platonic form. In Plato's system of

ethics, for every type of being there exists one 'perfect' example, the ideal form, which exists in some non-physical realm and is constructed by God. Goodness of an individual is defined in terms of how close it is to the ideal form. So the goodness of an individual of a species is measured by how close it is to the maximum of the breeding probability distribution (and on how close the physical attributes of the creature match those which would normally follow from the ideal genome; so if it is mutilated by disease, damage or some other factor it is less good than an individual that is not so disordered). For rational animals, the probability of the species passing on the genes does not only include the variation of the genetic code of the species, but also the character, virtues and choices of the individual members of the society[14]. This is the measure of moral goodness in this system: how close is the character of a person to that ideal character who populates the society that is best equipped to beget and successfully raise children. Some cultures are better equipped for survival across the generations than others, and the common virtues of the individuals within the most successful possible of societies would be the starting point of any ethical inquiry.

I should point out that what I have sketched out here (in a very incomplete form) is not the entirety of the system of ethics which I hold to; although it is a part. Firstly, I prefer a more Aristotelian approach than Platonic; secondly, while 'Be fruitful and multiply and fill the earth and subdue it' is important principle derived from natural law, it is not the only principle; and the other principles combine to reduce the worst excesses (for example eugenics) which may be inferred from the above rules. However, I feel that if one is to reject Aristotelian classical theism and the consequent natural law ethics and adopt atheism, then the above Platonic considerations provide the best possibility for deriving an objective moral code. (Objectivity is necessary if an ethic system can be universally applied; following Hume it is also denied by most modern systems of ethics which all boil down to a subjective choice. One of the reasons I have a very low opinion of modern moralists is that they simultaneously believe that it is impossible to derive 'ought from is', and also that their therefore arbitrary and inconsistent opinions on what ought to be done and believed are self-evidently true and should be forced on society.)

Note that I have only discussed whether evolution by natural selection is consistent with classical theism; and I believe that it is (although classical theism does not require it). I have not discussed whether it is consistent with Christian (or any other religion's) revelation (for example, the bible): that is a separate question (with several challenges to be overcome), which is best discussed elsewhere.

[14] I measure goodness in terms of the society's (and ultimately the species') fitness to breed and correctly raise the next generation rather than an individual by asserting as a second premise the principle that moral qualities should be universal; and apply equally to all members of the species rather than have different rules for different individuals.

CHAPTER 10

Quantum mechanics and Field Theory

10.1. Introduction

The state notation I used in the previous chapter is particularly well suited for quantum mechanics (indeed, I took it from Dirac's notation for quantum mechanics). But before discussing quantum physics, I want to help introduce the notation by applying it to classical mechanics.

$|X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle$ represents the state of a particle of type X at location $\vec{\mathbf{x}}$ and with momentum $\vec{\mathbf{p}}$. We use this notation in our equations to represent not the particle but a location and momentum which a particle could have. These state vectors form an unchanging canvas upon which physics takes place. In one moment of time, a particle will be in that state; in the next it will move onto a different state. The states remain constant; what changes is which state is occupied by a given particle. In the original conception of mechanistic physics, the location and momentum together describe everything there is to know about the particle; with quantities such as rest mass and electric charge specified by the type X. In this picture, the states are complete; to say that a particle is in state $|X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle$ is to describe everything relevant to physics about the particle; equally for every existent particle, there will be one state vector which can be used to describe it. All we need to do connect each particle at each given moment in time to one particular state vector. Obviously, if the particles have additional qualities (such as spin), then we can easily extend the state notation to also list these alongside location and momentum; however, for this discussion, and to keep things simpler, I will restrict myself to just classifying the states by location and momentum.

The next thing we need is a symbol to represent the actual state of the universe, which I denote as $|\Psi(t)\rangle$. This tells us which particles are present or actual at a time t. We choose to define the notation in such a way that $|\Psi(t)\rangle$ can be written as the sum of the those $|X\rangle$ which represent actual particles at a given time. For example, if all the universe contains at time t_0 is one particle of type X with momentum $\vec{\mathbf{p}}_1$ and location $\vec{\mathbf{x}}_1$ and one particle of type Y with momentum $\vec{\mathbf{p}}_2$ and location $\vec{\mathbf{x}}_2$, then we would write $|\Psi(t_0)\rangle = |X(\vec{\mathbf{x}}_1,\vec{\mathbf{p}}_1)\rangle + |Y(\vec{\mathbf{x}}_2,\vec{\mathbf{p}}_2)\rangle$. Why write $|\Psi\rangle$ as the sum of two states, and not a different formula? Firstly, just adding the two actual states together is the easiest way we can express the concept that $|\Psi\rangle$ contains those states; secondly since this serves as the definition we have a great deal of freedom in how we represent the physical states, so we might as well choose the simplest notation; and thirdly the framework I will develop here can be used to predict experimental results and, as far as they have been tested, it does so to a very high precision.

Note that the addition sign in the expression $|X(\vec{\mathbf{x}}_1,\vec{\mathbf{p}}_1)\rangle + |Y(\vec{\mathbf{x}}_2,\vec{\mathbf{p}}_2)\rangle$ is used analogously to its use a symbol representing the addition of two numbers. I am not saying that we are adding together two states as we might add together 5 and 3 to get 8. States are wholly different types of things than numbers. Rather, the statement means that the particle is either in state $|X\rangle$ or in state $|Y\rangle$. The addition sign is used because a) we need to use some symbol to represent this; and b) the operator we need has the same basic properties of addition such as commutativity, associativity, identity and inverse, and the way it relates to another operation analogous to multiplication (assigning the likelihoods that the particle is in a particular state). We can thus use the symbol as though it were truly an addition

sign, as we would in any linear or vector algebra. Since the operator obeys the same rules as the familiar addition operator, it is convenient to use the same symbol to avoid having to invent something new and unfamiliar.

We define another quantity $\langle X(\vec{\mathbf{x}},\vec{\mathbf{p}})|$, which serves as a query: is there a particle of type X at location $\vec{\mathbf{x}}$ and momentum $\vec{\mathbf{p}}$ present? Thus for a many particle state vector $|\Psi\rangle$, which represents the whole physical system being studied, $\langle X(\vec{\mathbf{x}},\vec{\mathbf{p}})|\Psi\rangle$ is 1 if the state contains a particle of type X with the specified energy and momenta and 0 if it doesn't. Thus $\langle X(\vec{\mathbf{x}}',\vec{\mathbf{p}}')|X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle$ is zero unless $\vec{\mathbf{x}}'=\vec{\mathbf{x}}$ and $\vec{\mathbf{p}}'=\vec{\mathbf{p}}$, when it is one.

What we are particularly interested in is how $|\Psi\rangle$ changes in time.

Rather than keeping track of multiple particles, let us start by just considering a single particle, represented by the symbol $|\psi(t)\rangle$. At any given moment in time, this will consist of one of the $|X\rangle$ states, and which of these states it is will change from one moment to another. Thus in classical mechanics we can write,

$$|\psi(t)\rangle = \int d^3x d^3p\, c_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}\, |X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle \tag{310.1}$$

where the coefficient $c_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}$ is 1 (or, more precisely, a Dirac δ-function whose integral over an infinitesimal volume containing that point gives 1) if that particle exists in the state at time t and 0 otherwise. So what we are doing here is taking every possible single particle state, multiplying it by 1 if it exists and 0 if it doesn't, and summing up over all of them. The final result will be the sum over all states which exist; this is just a slightly more cumbersome (but ultimately more useful) way of saying that $|\psi\rangle$ is the sum over the states of all existent particles, which, in classical mechanics, represents the state of the whole system. For a single particle state, only a single particular $c_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}$ will be one at a given moment in time. This c thus describes the dynamics of the system, and how the particle moves from one state to another over time. If there are two particles present, two values of c will be one and the rest zero, or, if the two particles happen to have precisely the same location and momentum, there will be one non-zero c with value $\sqrt{2}$. Why $\sqrt{2}$ rather than 2? The reason is buried in the mathematical structure described below.

The dynamics is then described by an operator $\hat{T}$. This encodes Newton's equations of motion or their relativistic equivalents. Thus for a particle in state $|\psi\rangle$ we write

$$|\psi(t+\delta t)\rangle = (1+i\delta t\hat{T})\,|\psi(t)\rangle\,. \tag{310.2}$$

$1+i\delta t\hat{T}$ is basically a set of rules saying that if you input a state with location $\vec{\mathbf{x}}_0$ and momentum $\vec{\mathbf{p}}_0$ you output a state with location $\vec{\mathbf{x}}_1$ and momentum $\vec{\mathbf{p}}_1$ as determined by the equations of motion. i is a symbol denoting the square root of minus one, and I have put this in here for reasons I will explain later. For example, in Newtonian mechanics, in a tiny time step δt, the location of the particle is changed by δt multiplied by the velocity, which is the momentum divided by the mass, while the momentum is changed by δt times the force (this follows from Hamilton's equations). So if particle X has mass m and is subject to a force $\vec{\mathbf{F}}$, then $(1+i\delta t\hat{T})c_{\vec{\mathbf{x}},\vec{\mathbf{p}},t} = c_{\vec{\mathbf{x}}+\frac{\vec{\mathbf{p}}}{m}\delta t,\vec{\mathbf{p}}+\vec{\mathbf{F}}\delta t,t+\delta t}$. This equation just means what I expressed in words in the previous sentence. The 1 in $1+i\delta t\hat{T}$ will later be used to remind us that the evolution is continuous, i.e. that the state at time $t+\delta t$ is almost the same as the state at time t, with only a small correction. The output state depends on the input state and the time interval, but nothing else. This is just the requirement that the evolution of a physical system just depends on the current state of the system and not anything else from its past history.

We can make measurements on $|\psi\rangle$ using other operators. For example, the operator

$$\hat{\vec{\mathbf{p}}} = \int d^3x d^3p\, |X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle\, \vec{\mathbf{p}}\, \langle X(\vec{\mathbf{x}},\vec{\mathbf{p}})| \tag{310.3}$$

returns the total momentum of the system multiplied by the original state. We can see this by looking at how this system can be applied to an input vector. Suppose that the

input vector is in one particular single-particle state $|X(\vec{\mathbf{x}}_1, \vec{\mathbf{p}}_1)\rangle$. In this case, the operator $\langle X(\vec{\mathbf{x}}, \vec{\mathbf{p}})|$ will return 1 if $\vec{\mathbf{x}}_1 = \vec{\mathbf{x}}$ and $\vec{\mathbf{p}}_1 = \vec{\mathbf{p}}$, and zero otherwise. Since we sum over every possible $\vec{\mathbf{x}}$ and $\vec{\mathbf{p}}$, it will pick out specifically that one state we are in. This then gives the output $\vec{\mathbf{p}}\,|X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle$ for the particular state where $\langle X(\vec{\mathbf{x}}, \vec{\mathbf{p}})|\psi\rangle = 1$, which is no more than $\vec{\mathbf{p}}_1\,|\psi\rangle$. If we just want the momentum of the a single particle represented by $|\psi\rangle$, then it is given by $\langle\psi|\,\hat{\vec{\mathbf{p}}}\,|\psi\rangle$. If $|\Psi\rangle$ represents the sum of various different states (i.e. multiple particles), then $\langle\Psi|\,\vec{\mathbf{p}}\,|\Psi\rangle$ will give the total momentum of the system. Similarly, we can have a location operator

$$\hat{\vec{\mathbf{x}}} = \int d^3x d^3p\,|X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle\,\vec{\mathbf{x}}\,\langle X(\vec{\mathbf{x}}, \vec{\mathbf{p}})| \tag{311.1}$$

so that $\langle\Psi|\,\hat{\vec{\mathbf{x}}}\,|\Psi\rangle$ measures the sum of the locations of the particles described by $|\Psi\rangle$, and a number operator

$$\hat{N} = \int d^3x d^3p\,|X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle\,\langle X(\vec{\mathbf{x}}, \vec{\mathbf{p}})| \tag{311.2}$$

so that $\langle\Psi|\,\hat{N}\,|\Psi\rangle$ counts how many particles of type X are contained within the state $|\Psi\rangle$ (this is not the same as $\langle\Psi|\Psi\rangle$ because $|\Psi\rangle$ might contain particles of a different type). In terms of the coefficients c, $\langle\Psi|\,\hat{N}\,|\Psi\rangle$ gives the sum of the squares of the c values. This is why a two particle state is represented by $c = \sqrt{2}$ rather than $c = 2$. Equally, we can measure the energy, $\mathcal{E}$, by using the Energy operator,

$$\hat{\mathcal{E}} = \int d^3x \int d^3p\,|X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle\,\mathcal{E}(\vec{\mathbf{x}}, \vec{\mathbf{p}})\,\langle X(\vec{\mathbf{x}}, \vec{\mathbf{p}})|\,. \tag{311.3}$$

This way of doing classical mechanics, introducing a continuous set of states and constructing operators to make measurements, is a little bit odd and adds a slightly cumbersome notation, especially since thus far we have forced each coefficient c to only be able to take the values $0, 1, \sqrt{2}, \sqrt{3}, \ldots$, which makes calculus (which requires continuous functions) difficult. It is not natural in terms of the mechanistic picture. However, both Newtonian mechanics and relativistic mechanics can be constructed in this way.

We can also use this picture if we are uncertain about the precise location of the particles. Forcing c to be precisely 0 or 1 assumes that we know the precise details of the energy and momentum of the state. But what if we don't know it exactly? This is, after all, our real life situation due to experimental error[1]. This state notation can also deal with that situation. All we do is simply allow c to be something between 0 and 1 and redefine $|\Psi\rangle$ so that instead of representing the actual state of the universe it represents our knowledge of the state of the universe, with the c value representing (in some way) our uncertainty as to whether that state is actual. In practice we can also express c as a complex number (why restrict ourselves to just real numbers?[2]), which only requires one modification to our notation: when we construct a query state $\langle\Psi|$ from the physical state $|\Psi\rangle$ we will need to take the complex conjugate of c – this ensures that the operators used to compute the average values of observable quantities such as $\hat{\vec{\mathbf{x}}}$ and $\hat{\vec{\mathbf{p}}}$ still return real numbers. The only requirement we have is that using the number operator applied to a single particle state $|\psi\rangle$

[1]Although in reality it turns out to be more fundamental than this. Quantum mechanical uncertainty is due to more than just the limitations of experimental precision. It is built into the structure of the universe. I mention experimental error here simply as a motivation.

[2]Our ultimate goal is to find a notation which can represent reality. We do not know at this stage whether reality requires c to be complex, or whether we can restrict it to just real values. If we keep it real, then if in practice reality needs complex numbers, we will at some point reach a dead end, have to tear up all our work, and start again. If, on the other hand, we start by allowing c to be complex, but in practice reality only requires it to be real, then since the real numbers are just a subset of the complex numbers, then our notation will still apply to the real world. Our description will be more than we need, but it will cost us nothing to prune it back if necessary at the end of the calculation after comparing our results against experiment.

should show that there is precisely one particle somewhere in the universe at each moment in time, which mathematically means that

$$\langle\psi|\left(\int d^3xd^3p\,|X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle\,\langle X(\vec{\mathbf{x}},\vec{\mathbf{p}})|\right)|\psi\rangle = \int d^3xd^3p c^*_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}c_{\vec{\mathbf{x}},\vec{\mathbf{p}},t} = 1, \tag{312.1}$$

where c^* indicates the complex conjugate. One of the features of complex conjugation is that a number times its complex conjugate is always positive and real, which makes sense for a count of the number of particles. The integral above is over all space and momentum, but we can also integrate over a specific volume or specific region of momentum only. For example, if we restrict ourselves to a volume V in space, (or we could likewise restrict ourselves to a specific range of momentum), then

$$\langle\psi|\left(\int_V d^3xd^3p\,|X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle\,\langle X(\vec{\mathbf{x}},\vec{\mathbf{p}})|\right)|\psi\rangle = \int_V d^3xd^3p\; c^*_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}c_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}, \tag{312.2}$$

lies between zero and one, and can be interpreted as a probability that the particle lies within that volume. This probability will describe the uncertainty we have over the precise location and momentum of the particle, which ultimately arises over the imprecision of our experimental measurements of the particle's location, momentum, energy and so on. So what we do is measure the distribution representing our knowledge of the particle's location and momentum given our initial measurement; construct a likelihood that corresponds to this probability distribution, use the equations of motion to evolve the likelihood over time, and consequently we can compute the probability that the particle will be in a particular location at some future time.

Why can we interpret $c^*_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}c_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}$ as a probability density? Firstly, it has all the properties of probability: it sums to one over all possible volumes, it is real and positive definite, and the probability that it is in the combined region $A \cup B$ is the probability that it is in region A plus the probability that it is in region B, as long as A and B don't overlap. Secondly, like classical probability, it represents a count over states: for example, if we assigned each state an equal likelihood, then the probability would represent the number of states in that region over the total number of states: this works out in the same way. Thirdly, we haven't yet defined what c physically represents, so we might as well define it so that c^*c is a probability: we are perfectly free to do so given that this quantity satisfies the properties listed above. c is related to the chance that the process occurs, but is not a probability in the strict mathematical sense (since it doesn't satisfy the defining properties of a probability), so I will refer to it as a *likelihood*.

I will make a few definitions here: the particle described by $|\psi\rangle$ is in a precise state (with respect to $|X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle$) if $|\langle X(\vec{\mathbf{x}},\vec{\mathbf{p}})|\psi\rangle| = 1$ for some particular $|X\rangle$ and consequently zero for all other possible state vectors, and a linear supposition of states if that is not the case. If it is in a precise state, then we can predict precisely what the value of the observable will be when we take a measurement; if it is in a linear supposition of states then all we can do is assign a probabilities (given by $c^*_{\vec{\mathbf{x}},\vec{\mathbf{p}}}c_{\vec{\mathbf{x}},\vec{\mathbf{p}}}$) that we will measure the particle to have location $\vec{\mathbf{x}}$ and momentum $\vec{\mathbf{p}}$.

What this interpretation means is that if we measure the initial condition at time 0, ($|\Psi(0)\rangle$), not as a definite state but as a distribution over a number of states, we can then apply the time evolution operator, $\hat{T}$, to get the state at a new time ($|\Psi(t)\rangle$),

$$\begin{aligned}|\Psi(0)\rangle &= \int d^3xd^3p\; c_{\vec{\mathbf{x}},\vec{\mathbf{p}},0}\,|X(\vec{\mathbf{x}},\vec{\mathbf{p}})\rangle \\ |\Psi(t)\rangle &= \lim_{\delta t\to 0} T[(1+i\delta t\hat{T})^{\frac{\Delta t}{\delta t}}]\,|\Psi(0)\rangle \\ &= T[e^{i\Delta t\hat{T}}]\,|\Psi(0)\rangle\,.\end{aligned} \tag{312.3}$$

and from this new state we can use the operators given above to compute the various observables: the mean value of the momentum, the mean value of the energy, its standard

deviation, and so on, using the new distribution. This is effectively the same as solving the equations of motion for each possible state, computing the observable, keeping track of all those results and then taking the appropriate weighted average at the end. Instead of doing each of these steps individually, this notation allows us to write down this whole process in one simple step. In the equation above, I have used a mathematical identity to write the product of the update steps as an exponential. The symbol T (without the hat) indicates time ordering, which means that we divide the exponential into small time segments, and then treat each one slice by slice. The symbol Δt indicates the duration we want to evolve the system to, our final time minus our initial time. This is, of course, $\Delta t = t - 0 = t$, but I prefer to write Δt rather than t because otherwise the symbol t would be put to too many uses: a symbol indicating the current time of the state $|\Psi\rangle$ and the total duration of interest. It happens that in this example these have the same numerical value, but in practice they are different things, and so we ought to use different symbols to represent them. The reason I wrote the i in the time update procedure is because we need to ensure that $\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|\psi(0)\rangle = 1$, to ensure that the number operator always says that we have one particle at each moment of time. This creates a condition on the operator $\hat{T}$. As I have written it, it means that $\hat{T}$ needs to be a Hermitian operator[3], which is good, because all the operators representing observable qualities are also Hermitian: they have to be to ensure that the average values, which correspond to our measurements, are real numbers.

We can, of course, also consider the probability that something remains in the same state over time. One of the basic rules of this system is that we can only compare two states if they represent particles at the same time. This is physically motivated: we can only compare things directly if we see them simultaneously. If they are at different times, we can only compare a description of one of them with the other, i.e. an indirect comparison. But that's OK: we have just shown that we can use the time evolution operator to translate states from one time to another. We have a given state at time 0, $|\psi(0)\rangle$. We need an operation asking whether the physical system is in that state: this is $|\psi(0)\rangle\langle\psi(0)|$. Thus the probability that a particle at time t is in the same state as at time 0 is given by

$$\langle\psi(t)|\psi(0)\rangle\langle\psi(0)|\psi(t)\rangle = \langle\psi(0)|\,T^{\dagger}[e^{-i\Delta t\hat{T}}]\,|\psi(0)\rangle\,\langle\psi(0)|\,T[e^{i\Delta t\hat{T}}]\,|\psi(0)\rangle\,, \tag{313.1}$$

where $T^{\dagger}$ means that we take the time ordering in the opposite direction, with the evolution operator computed at later times listed first (on the right) and earlier times last (on the left). This means that we can use the time evolution operator to construct probabilistic comparisons between states at different times.

We actually know what the time evolution operator is: it's just the differential operator $-i\frac{\partial}{\partial t}$, the operator that measures changes over time. The reason we can say this is from the Taylor series expansion[4] of $\psi(t)$, which happens (for small Δt) to give the same result that we get by expanding the exponential. But there is something strange about this. Because we also required that $\hat{T}$ represented the effect of the equations of motion acting on $|\psi\rangle$. Thus we can see that $\frac{\partial}{\partial t}|\psi\rangle = ???\,|\psi\rangle$, where ??? represents the operator that describes the equations of motion.

---

[3]A Hermitian operator is the complex analogue of a symmetric operator, i.e. if $\hat{H}|\psi'\rangle = |\psi\rangle$ then $\langle\psi'|H = \langle\psi|$; this is not a general feature of matrices but only one specific class of them; one could absorb the factor of i into $\hat{T}$, but it is conventional to write it this way with $\hat{T}$ Hermitian.

[4]One of the tricks used in mathematics is to find polynomials, $a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots$, which closely resemble a more complicated function near a point such as $x = 0$. The reason we do this is because polynomials are generally relatively easy to manipulate, while other functions aren't. The Taylor series expansion is a particularly important form of polynomial expansion where the value of the function at the specified point, its gradient, the rate of change of its gradient, the rate of change of the rate of change of the gradient, and so on to infinity, all precisely match between the original function and the polynomial. For most functions this means that the infinite order Taylor series expansion is indistinguishable from the original function, at least within a certain radius of convergence. Of course, what I am describing here is the expansion of a function of an operator rather than the expansion of a function of a variable, but the same principle applies.

This suggests that we can use differential operators as observables as well. Well, why not? So we can introduce another observable operator $\hat{\vec{\mathbf{P}}}$, which uses the differential operator $\vec{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$, which measures changes over spatial location:

$$\hat{\vec{\mathbf{P}}} = \int d^3x d^3p \ |X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle \, i\hbar\vec{\nabla} \, \langle X(\vec{\mathbf{x}}, \vec{\mathbf{p}})| \, . \tag{314.1}$$

Now, it is not clear at the moment what this represents. $\hbar$ is just a constant, which I have put in there because I can. I will give some physical meaning to it later. The $\vec{\nabla}$ operator acts on everything to its right. If we sandwich this between a couple of $|\psi\rangle$ states, we obtain

$$\langle\psi| \hat{\vec{\mathbf{P}}} |\psi\rangle = \int d^3x d^3p c^*_{\vec{\mathbf{x}},\vec{\mathbf{p}},t} i\hbar\vec{\nabla} c_{\vec{\mathbf{x}},\vec{\mathbf{p}},t}. \tag{314.2}$$

This measures how the probability amplitude c varies with location. Is this interesting? Well, the variation of c in time is certainly interesting, and from special relativity we know that time and space are mixed together, so this implies that it ought to be interesting in some respect. But what it represents is still not yet clear. However, one important thought is that the differential operator is the generator of translations: $e^{\vec{\alpha}\cdot\vec{\nabla}} f(\vec{\mathbf{x}}) = f(\vec{\mathbf{x}} + \vec{\alpha})$ for any function f and constant vector $\vec{\alpha}$. Obviously in classical physics, the velocity plays the role of generating translation, since having a velocity just means you are moving from one location to another. Thus this operator is in some way related to velocity, or perhaps momentum.

The interesting thing is what happens if we use both the operators $\hat{\vec{\mathbf{P}}}$ and $\hat{\vec{\mathbf{x}}}$. The result we get depends on the order in which we apply the operators. So the two measurements,

$$\begin{aligned} &\langle\psi| \hat{\vec{\mathbf{P}}} \cdot \hat{\vec{\mathbf{x}}} |\psi\rangle \\ &\langle\psi| \hat{\vec{\mathbf{x}}} \cdot \hat{\vec{\mathbf{P}}} |\psi\rangle \end{aligned} \tag{314.3}$$

do not give the same result. The reason is that in the first measurement we have to differentiate both the $\vec{\mathbf{x}}$ contained within the operator $\hat{\vec{\mathbf{x}}}$ and the $\vec{\mathbf{x}}$ contained within ψ, while in the second term we just differentiate the $\vec{\mathbf{x}}$ that is contained within ψ. The difference between these two operators can be measured using the commutator,

$$[\hat{P}_k, \hat{x}_j] = \hat{P}_k\hat{x}_j - \hat{x}_j\hat{P}_k = i\hbar\delta_{kj}, \tag{314.4}$$

where the indices indicate the component of the vector, and δ_{kj} is 1 when $k = j$ and 0 otherwise. Similarly, if we made this notation four dimensional and introduced a time operator, it would not commute with the time evolution operator.

One of the features of Hermitian operators is that they have eigenstates, which means that the operator acting on the eigenstate just gives the eigenstate multiplied by some number, the eigenvalue (similar to the eigenvectors discussed in the previous chapter; but $|\psi_*\rangle$ is a state rather than vectors so I call it an eigenstates). This includes the time evolution operator. So there is some state $|\psi_*\rangle$ such that

$$\hat{T} |\psi_*\rangle = \frac{E}{\hbar} |\psi_*\rangle \, , \tag{314.5}$$

for some number which I have called $\frac{E}{\hbar}$. This means that

$$e^{i\Delta t\hat{T}} |\psi_*\rangle = e^{i\Delta t\frac{E}{\hbar}} |\psi_*\rangle \, . \tag{314.6}$$

Now if we measure the probability that $|\psi_*\rangle$ remains in the same state after a period of time, we find

$$\langle\psi_*(t)|\psi_*(0)\rangle\langle\psi_*(0)|\psi_*(t)\rangle = \langle\psi_*(0)| e^{-i\Delta t\frac{E}{\hbar}} |\psi_*(0)\rangle \, \langle\psi_*(0)| e^{i\Delta t\frac{E}{\hbar}} |\psi_*(0)\rangle = 1 \tag{314.7}$$

In other words each state $|\psi_*\rangle$ (and there are many of them, corresponding to each eigenstate of $\hat{T}$) is unchanging in time. It is perfectly stable. This isn't going to be the case for a state of definite location and momentum $|X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle$, but it tells us that some superposition of those

states is stable over time. And this is, of course, important for us to know, since we are interested in studying physical objects which, in the absence of interactions, remain stable or at the very least meta-stable in time (aside from changes in location due to inertia). These would correspond to the things we observe and can study in physics (or at least the building blocks of such objects).

What is this stable state? In practice, the simultaneous eigenstate of both the differential operator $\hat{\vec{\mathbf{P}}}$ and something related to the time evolution operator which I will call $\hat{H}$, $\hat{H} = |X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle (-i\hbar\hat{T}) \langle X(\vec{\mathbf{x}}, \vec{\mathbf{p}})|$, is given by

$$|\psi_*\rangle = \int d^3x' d^3p' e^{-i(\vec{\mathbf{k}}\cdot\vec{\mathbf{x}}' - Et')/\hbar} |X(\vec{\mathbf{x}}', \vec{\mathbf{p}}')\rangle \tag{315.1}$$

That expression we have in the place of c is just the solution to the wave equation. So this implies that the stationary state of the time differential operator and the space differential operator together is just a wave. Now, for the waves we know about in classical mechanics, the energy it carries is proportional to its frequency, here $E/\hbar$, while the momentum is proportional to the vector I have written as $\vec{\mathbf{k}}/\hbar$ (this is confirmed experimentally for various types of waves: basically, if each wavefront carries the same amount of energy – has the same amplitude – then the faster the waves travel, the more wave-fronts hit a target within a given time, and the more energy is transferred to the target). So this means that, if we choose $\hbar$ to be the right number, we can associate E with an energy and $\vec{\mathbf{k}}$ with a momentum. If we apply the operator $\hat{\vec{\mathbf{P}}}$ to this wave solution, we find

$$\langle\psi_*| \hat{\vec{\mathbf{P}}} |\psi_*\rangle = \vec{\mathbf{k}}. \tag{315.2}$$

So this measures the momentum of the wave. Thus $\hat{\vec{\mathbf{P}}}$ is a momentum operator. Similarly, $-i\hbar\frac{\partial}{\partial t}$ would represent an energy operator, at least for these stationary wave states.

10.2. Quantum Mechanics

10.2.1. The Theoretical Structure of Quantum Mechanics

We have almost reached quantum mechanics. There are only a few more steps to go (Townsend, 2012; Dirac, 1964; Binney and Skinner, 2013; Griffiths, 2016). We wrote $|\Psi\rangle$ as a superposition of states, and stated that this was necessary because of the lack of experimental precision. The first step to get to quantum mechanics is to remove that 'because of the lack of experimental precision'. It is simply necessary. Nature just seems to work that way. Physical states have a certain degree of fuzziness inherent within them. The second step that we need to take is that these stationary wave states have an important physical meaning. It is these states, not $|X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle$, that represent both what we usually think of as particles (such as electrons and protons) and what what we think of as waves (such as the light carrying particles, called photons). This means that $\hat{\vec{\mathbf{P}}}$ is not just an operator that happens to measure the momentum for these waves, it *is* the momentum operator, both for these waves and everything else. $\hat{H}$ is not just the operator that happens to measure the energy for these waves, it *is* the energy operator, both for these waves and everything else.

But this leads to something else. We expressed the basis state as $|X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle$, as a function of both location and momentum, as demanded by classical mechanics. But we have just stated that the momentum operator is just the differential with respect to $\vec{\mathbf{x}}$. This makes the variable $\vec{\mathbf{p}}$ redundant. In classical mechanics, the location and momentum were independent of each other, so we had to provide both of them. In quantum mechanics, they depend on each other, so if you have the variation of the probability distribution in one then you can compute the other. We don't need to specify the momentum; we can derive it from the variation of $|X(\vec{\mathbf{x}}, \vec{\mathbf{p}})\rangle$ over location. So let's just get rid of all dependence of $|X\rangle$ on $\vec{\mathbf{p}}$. The basis state can just be written as $|X(\vec{\mathbf{x}})\rangle$, with no need to specify the momentum. Equally,

the operators lose that integral over momentum, so the number operator is just,

$$\hat{N} = \int d^3x \, |X(\vec{\mathbf{x}})\rangle \langle X(\vec{\mathbf{x}})| \,, \tag{316.1}$$

and the Energy operator just

$$\hat{H} = \int d^3x \, |X(\vec{\mathbf{x}})\rangle \, (-i\hbar \frac{\partial}{\partial t}) \, \langle X(\vec{\mathbf{x}})| \,. \tag{316.2}$$

This link between the differential operators and energy and momentum informs us about one of the earlier issues we had. We needed that $\hat{T}$ encoded both the equations of motion and the differential operator with respect to time. This implied that there had to be some equation linking the two. We now know what that equation is. We know that in Newtonian mechanics $E = \frac{(\vec{\mathbf{p}}-Q\vec{\mathbf{A}})^2}{2m} + QV(\vec{\mathbf{x}})$, for some external electrostatic potential V, magnetic potential $\vec{\mathbf{A}}$, electric charge Q, and particle mass m. Replacing E and $\vec{\mathbf{p}}$ with the relevant operators means that

$$-i\hbar \frac{\partial}{\partial t} |\psi\rangle = \left(\frac{1}{2m} (i\hbar \vec{\nabla} - Q\vec{\mathbf{A}})^2 + QV(\vec{\mathbf{x}}) \right) |\psi\rangle \,. \tag{316.3}$$

This is known as the Schrödinger equation. Similarly, in relativistic mechanics, we have $(E - QV)^2 = (\vec{\mathbf{p}} - Q\vec{\mathbf{A}})^2 c^2 + m^2 c^4$, which converts to

$$(i\hbar\partial_\mu - QA_\mu)(i\hbar\partial^\mu - QA^\mu) \, |\psi\rangle = m^2 \, |\psi\rangle \,, \tag{316.4}$$

where ∂_μ is the differential four-vector, $(\frac{\partial}{\partial t}, \vec{\nabla})$ (I am using units where the speed of light c has the value 1. I am free to do this, since I am free to choose whatever units I like to measure distances and durations; so, for example, I can measure time in seconds and distances in light seconds, and then in those units the speed of light will be 1). This equation is named after Klein and Gordon. These equations describe the time evolution of a particle.

These equations, especially the Schrödinger equation, are of vast importance: physics undergraduates spend what seems to be almost all their time learning how to solve it for various different choices of potential $V(\vec{\mathbf{x}})$, and using the solutions to make physical predictions. For example, the Schrödinger equation contains within it a very accurate (though not quite perfectly accurate – that needs quantum field theory) description of the atom: know how to solve the Schrödinger equation and you have atomic physics; if you have atomic physics, then you have chemistry. It tells us what the stable atomic states (the energy eigenstates of the atom) are; more importantly, it tells us that the atom is stable, and won't collapse when the electrons radiate into the nucleus, which would happen in a classical mechanics description of the atom, leaving just inert blobs of matter. Almost everything in our world around us can be described using equation (316.3); if the speed of light were infinite (which in particular means that an infinite amount of energy would be needed to convert into even the smallest non-zero mass) and there were no gravity, it would, when combined with statistical mechanics, give us a perfect description of nature; as it is, its solutions provide a very good first approximation to almost every physical system. If the universe was classical rather than quantum, then almost nothing in our world, even on a macroscopic scale where classical mechanics is a very good approximation, would exist as we observe it: without atoms, there would be no chemistry, and consequently no biology, no rocks or planets or stars: no structure in the universe at all. So if you want to learn how to make predictions from physics, then understanding the Schrödinger equation is one of the most important steps along the way. However, I intend to go in a different direction, so I will leave the Schrödinger equation unsolved and pass on to my next topic.

The next step to get to quantum mechanics is to say that when we measure a quantity represented by the operator $\hat{Y}$, then the eigenstates of $\hat{Y}$ play an important role: the eigenvalues represent the only numbers we can get from the measurement. If we measure the quantity, then we will get one of the eigenvalues. Indeed, by making the measurement,

we force $|\psi\rangle$ into one of the eigenstates of $\hat{Y}$. I'll call this eigenstate $|Y_n\rangle$ and the eigenvalue Y_n, so we have the equation $\hat{Y}\,|Y_n\rangle = Y_n\,|Y_n\rangle$ as the definition of the possible values of Y_n and $|Y_n\rangle$. This process of forcing the particle into an eigenstate of the measurement operator is known as wavefunction collapse. It is not determinate (unless the particle was originally in an exact eigenstate of $\hat{Y}$), which means that you cannot predict it in advance. What is the probability that you get a particular result? We already know that. The operator asking 'is this state in state $|Y_n\rangle$?' is $|Y_n\rangle\,\langle Y_n|$, by analogy with our discussion above. Thus the probability that we will measure the value Y_n is just $\langle\psi|Y_n\rangle\langle Y_n|\psi\rangle$. But the point is that we are now talking about probabilities. The process of measurement in quantum mechanics is not deterministic, although the evolution of the one particle amplitude $|\psi\rangle$ in the absence of measurement is deterministic. This role played by measurement in turning a deterministic theory into an indeterminate theory is rather troubling and confusing, since it seems to suggest that we ourselves (or any conscious rational agent capable of making measurements) play an important role in the evolution of physics, which runs against the idea of the ideal experimenter as a impassive observer of what he is studying. But I don't want to address this question further at the moment, and certainly not until after I have introduced quantum field theory.

Quantum mechanics is not the most fundamental theory: that's quantum field theory, and a number of things change between the two. There is no point in worrying about the philosophy of quantum mechanics, since it doesn't represent the real world. We need to think about the philosophy of field theory instead. In terms of physics, quantum mechanics is an important bridge between classical mechanics and quantum field theory, which is why I am discussing it. But there is no point yet trying to wrap your head around all this and trying to develop a viable philosophy: you will just give yourself a headache for no reason because quantum mechanics doesn't correspond to reality, so whatever philosophy you come up with likely won't correspond to reality either[5] Keep your philosophising, and your headaches, until later, and the discussion of quantum field theory. Quantum mechanics is a compromise position between the mechanistic classical physics and the Aristotelian quantum field theory, built on some premises from one philosophy and some from the other. But, with apologies to Hegel, the fusion between a contradictory thesis and antithesis is not a synthesis but a self-contradictory nonsense. Such it is with the philosophy of quantum mechanics, with the difficulties in understanding it arising from this contradictory nature. For the moment, just try to follow the mathematics (which is possibly a headache in itself) and don't worry too much about what it means at the philosophical level.

I mentioned that momentum was redundant as a descriptor of the basic states $|X\rangle$. This isn't actually quite true; it is true that we only need one variable, but we have a lot of freedom in determining what that variable is. Mathematically the key properties we need of the basis states $|X\rangle$ are that they form a complete set, which means that every possible physical state can be expressed as a linear combination (or superposition) of the $|X\rangle$ states, and secondly that they are what is known as orthonormal, i.e.

$$\langle X(\vec{\mathbf{x}}_1)|X(\vec{\mathbf{x}}_2)\rangle = \delta(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2), \tag{317.1}$$

[5]Some interpretations of quantum mechanics carry over to field theory with little modification, such as the standard Copenhagen interpretation. But there are other interpretations of quantum mechanics, such as the De-Broglie/Bohm interpretation, which, despite the best efforts of some researchers, are generally considered to be incompatible with quantum field theory. So the question comes why we should suppose that every possible interpretation of quantum field theory, including maybe even the true interpretation, will have a quantum mechanical analogue? By concentrating on quantum mechanics, there is a chance that the philosopher of science might be missing something important. Or they might not be. We don't know at this point of the argument. But you don't lose anything by using field theory as the starting point of your philosophy, while you might lose something if you base it on quantum mechanics.

where δ, the Dirac δ-function, indicates a function that is zero whenever $\vec{\mathbf{x}}_1 \neq \vec{\mathbf{x}}_2$ and infinite when $\vec{\mathbf{x}}_1 = \vec{\mathbf{x}}_2$, but in such a way that the integral over the δ function gives 1[6]. The $|X(\vec{\mathbf{x}})\rangle$ states are the eigenstates of the location operator $\hat{\vec{\mathbf{x}}}$: they represent, after all, the allowed locations of the particle, and in quantum mechanics the eigenvalues and eigenstates represent the allowed measured values and the allowed measured states. This actually guarantees our two needed properties of completeness and orthonormality, since these are well known properties of the eigenstates of any Hermitian operator, and $\hat{\mathbf{x}}$ is a Hermitian operator. But the phrasing *any* Hermitian operator should strike our interest. All we need for our mathematical notation of possible states to work are these two conditions. So in principle *any* Hermitian operator should do the job. We could use the eigenstates of the energy operator $\hat{H}$ as our underlying basis states, and express our particle state $|\psi\rangle$ as a superposition of these states. We could use the eigenstates $|P(\vec{\mathbf{p}})\rangle$ of the momentum operator $\hat{\vec{\mathbf{P}}}$ as our basis states. This suggests that there is a relationship between the $|X\rangle$ states and the $|P\rangle$ states, and here it is:

$$|P(\vec{\mathbf{p}})\rangle = \int \frac{d^3x}{\sqrt{2\pi\hbar}} e^{i\vec{\mathbf{p}}\cdot\vec{\mathbf{x}}/\hbar} |X(\vec{\mathbf{x}})\rangle$$
$$|X(\vec{\mathbf{x}})\rangle = \int \frac{d^3p}{\sqrt{2\pi\hbar}} e^{-i\vec{\mathbf{p}}\cdot\vec{\mathbf{x}}/\hbar} |P(\vec{\mathbf{p}})\rangle\,. \tag{318.1}$$

This means that in quantum mechanics it is possible to express the theory so that the momentum alone is sufficient to describe the allowed states of the system. In classical mechanics we needed both momentum and location to describe a state. In quantum mechanics we can (though we don't need to) just use the momentum. This is important for our overall thesis. The only major objection to scholastic metaphysics (i.e. the only valid objection that I couldn't solve with a relatively minor modification of scholastic metaphysics) from classical mechanics was that Newton's first law was only equivalent to the principle of causality if the state of the particle could be expressed only in terms of the momentum. But in classical mechanics, we need both momentum and location to define the state, so the principle of causality didn't work. Now we discover that one can express quantum mechanical states in terms of the momentum alone: there is no need for location. In other words, combine this with Newton's first law, and the objection to stochastic metaphysics disappears: the state of the particle can be expressed in terms of the momentum only, and the momentum (and thus the state of the particle) can only change under an external force. This is the scholastic principle of causality. Change is just reduced to a change of momentum[7], and momentum can only be changed if something external changes it. Of course, there is a drawback: we have removed this problem, and created another in its place. This is quantum mechanics, not classical mechanics, and Newton's first law might no longer be applicable. So we haven't solved any philosophical problem yet. But hold this thought in your mind: we will come back to it.

I should also go back to the commutation relation, equation (314.4). There is an important physical significance here.

$$[\hat{P}_k, \hat{x}_j] = i\hbar\delta_{kj}. \tag{318.2}$$

Going back to the mathematics, we know that operators which commute (the operators $\hat{A}$ and $\hat{B}$ commute if $[\hat{A}, \hat{B}] = 0$) share common eigenstates. If they don't commute, then they

[6]Imagine a rectangle of constant area, which is the base times the height. As you decrease the size of the base while keeping the area constant the height must increase; decrease the size of the base towards zero to represent a single point, and the height becomes infinite while the area remains the same. The Dirac δ-function is the limiting case of that family of rectangles as the base size reduces to zero.

[7]Actually, the state is not just a momentum state, but also depends on spin, colour and qualities such as that. However, this doesn't actually affect what I am saying. The key point is that we do not need both momentum and location to define a state.

don't share common eigenstates. What this means is that it cannot be in a simultaneous eigenstate of the location and momentum operators at the same time. The eigenstates represent the allowed physical states which correspond to what we observe when we take a measurement. This means that the particle cannot simultaneously be in a precise location state and a precise momentum state simultaneously; if it is in a precise location state it must be in a linear superposition of momentum states, and vice versa. If we measure the location of the particle, we throw it into an imprecise momentum state (or linear superposition of momentum states). If we measure the momentum, we throw the particle into an state of imprecise location. It is not only that we need not have the state of the particle described by both location and momentum, we cannot do so. This is the basis of Heisenberg's uncertainty principle: one cannot simultaneously precisely measure two observables which are represented by non-commuting operators.

Note, however, the role of $\hbar$ in this equation. If $\hbar$ were zero, then the momentum and location operators would commute; one could know location and momentum simultaneously. Indeed, we would have to specify both observables in order to express the full state of the system, which means that classical physics would still work. $\hbar$ is not zero, but it is exceptionally small in comparison to the sort of momentum and distance scales we encounter in our everyday lives. The smallness of $\hbar$ is the reason that classical physics works so well, and that it is non-zero the reason that classical physics is ultimately wrong. Classical mechanics is the $\hbar \to 0$ limit of the quantum theory. Once again, as in special relativity, we see that classical mechanics differs from the true theory by corrections controlled by a small parameter[8]. In the case of special relativity, this small parameter is the inverse speed of light c^{-1}. In the case of quantum mechanics, this small parameter is Planck's constant $\hbar$. There is, however, an important difference between the two cases. Special relativity doesn't change the underlying philosophy of classical mechanics; it is just a modification to the mathematical form of the laws of motion. Quantum physics represents a whole new paradigm. We cannot assume that any of the philosophical basis of classical mechanics will be valid in a quantum world. It might be, or it might not be: $[\hat{P}_k, \hat{x}_j] \neq 0$ and $[\hat{H}, \hat{t}] = -i\hbar \neq 0$ are so different from the classical picture that we cannot assume any of the underlying philosophy will still be valid.

The energy state of a free particle, which is just the wave state I mentioned above in equation (315.1), has continuous eigenvalues: E can take any real number. This isn't always true. Indeed, once we add a potential, in almost every case of physical interest, the eigenvalues of the time evolution operator as described by the Schrödinger or Klein-Gordon equations only permit a discrete set of possible eigenvalues. For example, let us consider an electron in a Coulomb electrostatic potential: this is the simplest possible description of the Hydrogen atom. It is not quite accurate, since the electrostatic potential is not the only 'force' acting on the electron (for example, there is an interaction between the magnetic field from the orbiting electron and the spin of the nucleus). However, the eigenstates of the energy operator, which represent the meta-stable states and therefore are the allowed values of the energy, are discrete: they come in powers of $E_n = 1/n^2$ for positive integer n. It was actually this observation that in part inspired the development of quantum mechanics. The angular momentum states of the electrons are also discrete, and for each energy level n there are n possible different angular momentum states. By finding the solutions to the Schrödinger equation for different n and angular momentum, we can compute the various probability distributions for finding the electron in a particular energy state in a particular

[8]'Small' here is a little misleading, since both the inverse speed of light and Planck's constant are dimensional quantities, i.e. they depend on our arbitrary choices of unit. One can choose the units so that the numbers representing them are actually quite large, in fact anything we want except zero and infinity. They only take on a particular value once we fix the length scales we use to make measurements. I mean by 'small' that they are small in comparison with the sort of energy, length and durations used when Newton's mechanics works to a good approximation.

location around the atom. Other atoms are more complicated, because one also has to include the interaction between electrons, but can also be solved.

These quantum mechanical eigenstates represent the stable states of an atom; place the electrons in the energy eigenstates and leave the atom alone, and they will remain in those states eternally (although the picture changes slightly when we move into field theory). They also represent the possible states of the atom; the electron must be in one of these eigenstates (although our knowledge of the electron's state might not be); one (albeit oversimplified) picture is that if the electron were sitting at an energy level between two eigenstates then it would be unstable, and thus its energy would change over time, until it moves into an energy eigenstate after which it will be fixed. In practice, in quantum field theory, the eigenstates represent meta-stable states, and there is a non-zero probability that the electron will spontaneously transition from one state to another, but the underlying structure of a sequence of states remains. These states are discrete, not continuous. For the free particle waves I discussed above, the particle could in principle adopt any energy value. The bound, or interacting, particle is, however, restricted to only certain discrete energy levels. The difference between these two situations is analogous to the difference between integers and real numbers. The idea that energy can only take certain discrete values, or equivalently that the particle can only be in a discrete set of states, is alien to classical physics, but is an unavoidable consequence of the Schrödinger equation.

Another important example is the harmonic oscillator[9], where the potential scales like $\frac{m}{2}\omega^2 x^2$: here the Energy operator eigenstates, and thus the allowed energy levels, scale like $E_n = (n + \frac{1}{2})\hbar\omega$ for positive integer n. In classical mechanics, the energy of the system can take any value; in quantum mechanics it can't.

In fact, the Harmonic oscillator in one space dimension and one time dimension is worth looking at in a little more detail, because the ideas I will present here will come up again in another context later. The energy operator is

$$\hat{H} = \frac{1}{2m}\hat{P}^2 + \frac{m}{2}\omega^2\hat{x}^2, \tag{320.1}$$

where ω is some coefficient used to control the strength of the interaction. $\frac{1}{2m}\hat{P}^2$ represents the kinetic energy of the particle, and $\frac{m}{2}\omega^2\hat{x}^2$ the potential energy; this equation states that the total energy is the sum of the kinetic and potential energy. We can introduce two operators, which I will call $\hat{a}$ and $\hat{a}^\dagger$, as follows

$$\begin{aligned}\hat{a} &= \frac{1}{\sqrt{2m}}\hat{P} + i\sqrt{\frac{m}{2}}\omega\hat{x} \\ \hat{a}^\dagger &= \frac{1}{\sqrt{2m}}\hat{P} - i\sqrt{\frac{m}{2}}\omega\hat{x}\end{aligned} \tag{320.2}$$

The Energy operator can then be written as

$$\hat{H} = \hat{a}^\dagger\hat{a} - i\frac{\omega}{2}[\hat{P}, \hat{x}], \tag{320.3}$$

where $[\hat{P}, \hat{x}]$ is the commutator introduced above, in equation (314.4). Thus,

$$\hat{H} = \hat{a}^\dagger\hat{a} + \frac{\hbar\omega}{2} \tag{320.4}$$

In fact, we can write the commutator between $\hat{a}$ and $\hat{a}^\dagger$ as

$$[\hat{a}, \hat{a}^\dagger] = \hbar\omega. \tag{320.5}$$

[9]This is a simple soluble model whose solutions can be used as a first approximation of more realistic systems – most systems which have a minimum in the potential can be written as a harmonic operator plus some small corrections, and we can find the solution to the harmonic operator and then systematically correct it. The harmonic oscillator is also useful for what it teaches us about quantum mechanics. For example, in classical mechanics, the frictionless pendulum is described by a harmonic oscillator to a very good approximation.

Now suppose that $|\psi_n\rangle$ is an energy eigenstate, i.e. $\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle$ with Energy eigenvalue E_n. Then we can see that $\hat{a}^\dagger|\psi_n\rangle$ is another energy eigenstate with eigenvalue $E_n + \hbar\omega$,

$$\begin{aligned}\hat{H}(\hat{a}^\dagger|\psi_n\rangle) =&(\hat{a}^\dagger\hat{a} + \frac{\hbar\omega}{2})\hat{a}^\dagger|\psi_n\rangle \\ =&\hat{a}^\dagger\hat{H}|\psi_n\rangle + \hat{a}^\dagger[\hat{a},\hat{a}^\dagger]|\psi_n\rangle = \\ =&(E_n + \hbar\omega)(\hat{a}^\dagger|\psi_n\rangle). \end{aligned} \tag{321.1}$$

Similarly, we can see that $\hat{a}|\psi_n\rangle$ is an energy eigenstate,

$$\hat{H}(\hat{a}|\psi_n\rangle) = (E_n - \hbar\omega)(\hat{a}|\psi_n\rangle). \tag{321.2}$$

$\hat{a}$ is known as the lowering operator and $\hat{a}^\dagger$ the raising operator because they have this effect of raising or lowering the energy levels. I should add that $\hat{a}^\dagger|\psi_n\rangle$ is not the next allowed energy state $|\psi_{n+1}\rangle$, but only proportional to it. It is easy to work out what that constant of proportionality is, but that goes beyond what I need to say here. The thing we need to get from this discussion is that in quantum theory the transition from one energy level to another is also described by certain operators, in this case products of either the raising operator $\hat{a}^\dagger$ or of the lowering operator $\hat{a}$. These operators also allow us to compute the dynamics of the system without having to directly solve the Schrödinger equation.

This seems to suggest that if E_n is an allowed energy level of the system, then there are other allowed energy levels separated from it by integer multiples of $\hbar\omega$ from plus infinity through to minus infinity. But this is a) troubling and b) not quite true. It is troubling because the energy operator $\hat{H}$ is made of two terms, one proportional to $\hat{P}^2$ and the other proportional to $\hat{x}^2$. These both only have positive eigenvalues, and since the sum of two positive numbers can only be a positive number, this suggests that the eigenvalues of $\hat{H}$ can also only be positive (of course, the minimum eigenvalue of $\hat{H}$ is not just the minimum eigenvalue of $\frac{m}{2}\omega^2\hat{x}^2$ plus the minimum eigenvalue of $\frac{1}{2m}\hat{P}^2$, because the two operators don't commute, and thus cannot have simultaneous eigenstates. This doesn't affect the argument that the smallest eigenvalue of the sum of these two operators must be greater than zero). Secondly, it is not true because it might be that for some energy state $|\psi_0\rangle$, $\hat{a}|\psi_0\rangle = 0$, in which case that state is the lowest we can get to via the lowering operator. As it happens, this state corresponds to $E_0 = \frac{1}{2}\hbar\omega$. Thus we can see that the allowed energy levels are $E_n = (\frac{1}{2} + n)\hbar\omega$ as promised above.

10.2.2. The Experimental Basis of Quantum Mechanics

My discussion of quantum mechanics so far has made one hugely important omission: experiment. I need to correct for this, for two reasons. Firstly because experimental testing is crucial for any study of physics. Secondly, for historical reasons. In reality, experiment drove the development of quantum mechanics entirely. For its three predecessors (in the order of the development of the ideas), Aristotle's physics, Newton's physics, and the theories of relativity (in terms of time, relativity was presented during the early development of quantum mechanics), philosophy had an important role in the development of the theory. In each case (even Aristotle) experiment played an important role, both to show there was something wrong with the prevailing view of physics and thus initiate thinking about the subject, to fine tune the final theory, and to test the theory. But also in each case the theory was developed along lines suggested by an overriding philosophical idea and as the culmination of various philosophical theories that were being developed simultaneously with the physics.

Quantum mechanics was different. Nobody saw it coming. There was no precursor to it in the history of non-scientific ideas; although its formalism borrows a great deal from nineteenth and twentieth century mathematics the application of that formalism was

wholly unexpected. The formalism I have just outlined developed because it was the only[10] way which people could reconcile theory with the new experiments. This remains true as quantum mechanics developed into quantum field theory: the development of the standard model of particle physics in the 1950s through to the 1970s was very much a case of theorists trying to keep up with the experimentalists. The difficulties we have today in combining quantum field theory with general relativity, and the lack of progress in particle physics since the 1980s[11] is very much a case of the theorists having caught up with experiment and waiting for something new to appear. Aside from the observation of neutrino masses, every experiment conducted since that time has slotted into the standard model.

So where does that leave the philosophers? Nowhere really. The initial response to quantum mechanics was to move from the objective and mechanistic modernism to more a subjective idealism (or post-modernism), basically a development from Kant, but this doesn't really lead anywhere because quantum mechanics (in its Copenhagen interpretation) may need the act of measurement to be incorporated into the theory, but it still remains objective because it doesn't matter who is doing the measuring. The Everett multi-world interpretation, where each act of measurement splits the universe into numerous branches, each corresponding to one possible outcome, is (I think) the only interpretation of quantum mechanics consistent with the determinism that underlies much of modern philosophy, but suffers from implausibility, from that it is not testable, and that it is very much ad-hoc: why on earth should the universe behave like this? But what if the universe is simply not, after all, deterministic: given complete information about the present, one cannot even in principle predict the future except probabilistically? This would mean re-examining almost all philosophy since Descartes. A new paradigm entirely is needed, and philosophers have had fifty years (since quantum field theory was established) to do this, but haven't. In other words, we physicists cannot be guided by philosophy, since almost all contemporary philosophers are still focusing on ideas which have been shown to be wrong. That means, of course, that we have to wait for our other source of inspiration, namely experiment, to overtake theory.

The electron was isolated in the late nineteenth century, and subsequent investigations studied how the positive and negative charges fit into the structure of the atom. Rutherford's scattering experiments indicated that there was a small core of positive charge at the centre of the atom, with electrons circling it. This lead to the 'planetary' model of the atom, with

[10]Again, I am being a little bit too simplistic by saying 'only' here. In one sense the formulation I have used is not the 'only' one; it is not even the formulation that students of physics typically first encounter. However, the different notations used to describe quantum mechanics are all mathematically equivalent. They are all saying the same thing, just written in different ways. It is in this sense of mathematical equivalence that I feel justified in using the word 'only'.

[11]Again, I should clarify. The final parts of the standard model were in place by the mid 1970s, with a few pieces of book-keeping settled towards the end of that decade. Since then, our understanding of what the standard model implies has grown considerably. We know things about the standard model, and have techniques to compute things in it, which go far beyond what the originators of the theory could have imagined. But this is just tinkering around the edges. The overall paradigm remains the same. An up-to-date 1980 textbook which contained everything known at that time wouldn't be enough to prepare a student for a research career in theoretical standard model physics today, but it would come pretty close. Obviously string theorists, advocates of supersymmetry, the Hartle-Hawking model, loop quantum gravity, causal dynamical triangulation, or others playing around with the numerous beyond the standard model ideas would reject my statement that there has been no significant progress in theoretical particle physics since that date. And it is true: we know a lot more about string theory today than we did in 1980. But we are no closer to knowing whether string theory (for example) is the correct fundamental theory of physics; it just remains speculation. If string theory should turn out to be correct (which is possible), then the last few decades would have turned out to have been a time of immense progress; if, on the other hand, string theory is wrong (which is also possible), then all that work would have just been a waste of time, baring its (admittedly quite beautiful) mathematical spin-offs. The case of cosmology – the other half of fundamental physics – is different. There have been huge leaps forward in cosmology, both experimental and theoretical, since the 1980s. And the same can be said for most other areas of physics, including experimental particle physics. It is just theoretical particle physics which has had the problem.

electrons orbiting the nucleus as the planets orbit the sun. However, this model didn't work for several reasons. It was known that the electrons emit electromagnetic radiation. This would lead to a loss of kinetic energy, and thus they would move into a closer orbit around the nucleus. However, in the classical picture, there was nothing to stop them from radiating away all of their kinetic energy and dropping into the nucleus itself. Planets have stable orbits in a vacuum, but this was more like trying to create a stable orbit in treacle, with the underlying electromagnetic field acting to suck away the motion of the electrons. The classical 'solar system' picture of the atom thus predicts that atoms ought to be unstable, which is clearly false. Secondly, it was known that energies of the radiation emitted by the atom came in discrete bands. In the classical picture, any energy emission was possible; that was not what was seen. Thus the solar system model is wrong. Indeed, atoms are nothing like that model, because they are governed by the rules of quantum mechanics rather than classical mechanics.

Focusing on hydrogen as the simplest atom, Bohr noted that the energy levels of the electrons in the atom were explicable if the electron's angular momentum was restricted to integer multiples of $\hbar$. This is actually one of the consequences of the solutions of the Schrödinger equation for the Coulomb potential, but that wasn't known yet. Instead it was apparent that energy was restricted to certain discrete packets – these were labelled *Quanta* after the Latin word for 'How much,' but which in physics came to represent the smallest allowed discrete package of Energy.

At the same time, another interesting effect was seen when the incidence of light stimulated a flow of electrons in certain materials. The energy given to the electron had to be delivered at a particular moment at a particular instant in time. Einstein (inspired by earlier work by Planck on black bodies) correctly proposed that this implied that light came in small packets – quanta – localised in space and time. This, however, is a feature of particles, not waves. These packets had an energy of $\hbar\omega$, where ω was the frequency of the light.

At a similar time, electrons were passed through two thin slits, and diffraction effects were observed - the characteristic behaviour of waves. But the strange thing is that the electrons still arrive one by one. If you fire one electron at a time through the two slits, you can see each electron hit the detector screen in a particular place. This is very much a particle like behaviour - fire a bullet through a hole in a metal plate and it will produce a single dent at the screen on the other side. Maybe it will hit the edges of the slit and its direction will be diverted, but it will not split up or seem to pass through both slits simultaneously. But as you continue the experiment with electrons, noting down where each electron hits, you will start to see the familiar diffraction pattern emerge. Close one of the slits, and the diffraction pattern disappears. Try to measure which slit the electron passes through, and the diffraction pattern disappears. In other words, there is an interference effect – but what is the electron interfering with? It can only be itself. But where each electron appears can't be predicted. It seems to be completely random. Thus both electrons (and other things traditionally called particles) and electromagnetic radiation have a seemingly dual nature sometimes they behave like classical particles and sometimes like classical waves.

Of course, the solution to this dilemma is that they are neither classical particles nor classical waves, but something different. This is something of a problem from the classical mechanistic world-view, where everything can be reduced to rigid lumps of matter, with a certain fixed extension, in motion. The classical picture is perfectly natural from the world we immediately observe, where quantum fuzziness is too small to notice, but it is wrong. Once again, the classical imagination, which underlies modern philosophy, has failed us. We still think of particles in modern physics, but we now longer mean the same thing by the word that people like Newton and Boyle did. Of course, what they are is another matter. Quantum mechanics doesn't tell us – but it does us that what they aren't: classical particles.

An analogy with optics immediately suggests that electrons should be described by a amplitude/intensity notation, since this is how interference effects can arise. The intensity

in this case has to represent a probability, allowing us to compute the distributions seen by individual electrons in diffraction. It is necessary to use probability because there is a gap in our knowledge – we can't predict the paths of the electrons. We thus need to introduce the idea that the intensity represents a probability density. This starts to lead us to the theoretical formulation of quantum mechanics. Inspired in part by Heisenberg's work, Schrödinger wrote down his equation, and a little later Dirac invented a notation similar to my presentation above, and showed that Heisenberg's matrix mechanics and Schrödinger's wave mechanics and his own state notation were mathematically equivalent. Schrödinger's equation solved the diffraction experiment, hydrogen atom, photoelectric effect and the other experimental observations. From this, the structure of the atom was (not yet completely accurately, but getting close) finally understood. The link was established between physics and chemistry. At the same time, quantum mechanics was merged with the nineteenth century statistical physics, leading to various other advances including a complete understanding of black bodies, and eventually the emergence of a complete and reasonably accurate (though again, the full theory still had to wait a little bit longer) theory of solids. There were still difficulties merging quantum mechanics with special relativity, but in the non-relativistic regime, progress was rapid.

Two new features emerged from experiments during this period in the early twentieth century. The first was the concept of spin, which, as mentioned earlier, is a feature which has similar properties to angular momentum, but isn't. It is described by the same mathematics as angular momentum, and interacts with magnetic fields in the same way. There is, however, an interesting property of spin, related to rotational symmetry. We can classify a particle based on how many times we have to 'rotate' it to return to the starting position. So, for example, spin 0 implies complete rotational symmetry (a spherical symmetry), spin 1 means that you have to 'rotate' through 360°to get back to an identical configuration, spin 2 'rotate' through 180° (a symmetry under a half-rotations) and so on. What is surprising is that many fundamental particles are spin $\frac{1}{2}$. If this were a matter of rotation and angular momentum, it would mean that you have 'rotate' the particle through two complete revolutions to get to an identical state. So if one takes a spin $\frac{1}{2}$ particle in a state $\left|\frac{1}{2}\right\rangle$ and R is the operator that gives a complete rotation through 360°, then $R\left|\frac{1}{2}\right\rangle = -\left|\frac{1}{2}\right\rangle$. Of course, spin isn't angular momentum, and this procedure isn't rotation - it just appears to be like that because of the similarity in the mathematics. But there was a distinction emerging between those particles with half integer spin, called Fermions, and those particles with integer spin, called Bosons.

This distinction is the second new feature of quantum mechanics. The main difference between the two is what happens when you interchange two particles of the same type (so these are both the same kind of particle – possibly both electrons – they do both have the same spin, but are in different energy eigenstates, for example). So let us say that we have two particles described by states $|A\rangle$ and $|B\rangle$. The combined state we describe as $|AB\rangle$. The question is what is the relationship between $|AB\rangle$ and $|BA\rangle$, which is the same state except that the two particles have been exchanged? The two states look similar, so one might think that there is no physical difference between them and therefore $|AB\rangle = |BA\rangle$. And this is the case with Bosons. However, for Fermions it is slightly different. Here we have $|AB\rangle = -|BA\rangle$. That minus sign leads to another property, namely that if we put two Fermions of the same type into the same energy eigenstate we get $|AA\rangle = -|AA\rangle$ which implies that $|AA\rangle = 0$. This means that you can't have two Fermions of the same kind with the same spin in the same energy eigenstate. This is known as the Pauli exclusion principle, and it is vitally important in almost all areas of physics and theoretical chemistry. Most of the fundamental particles – electrons, protons, neutrons, quarks, neutrinos, and so on, are spin $\frac{1}{2}$. This means that there can only be two electrons (one with positive spin, and the other negative spin) in the energy levels around each atom - which affects the ways that atoms can interact with each other, and ultimately gives us the periodic table where

atoms with a similar angular momentum eigenstates for their outermost electrons have similar properties, and those with different angular momentum eigenstates have different properties.

Almost every quantum mechanical system has various different discrete energy levels. The system can be excited from one level to a higher one by absorbing a quantum of light, or it can emit light to move down to a lower energy level. Thus there is not a single unique state which a quantum mechanical substance can exist in. All these states are described by the same Hamiltonian, so they need to be treated together, but the individual energy eigenstates need to be distinguished. At any given time, one of these eigenstates will be occupied, and the others empty.

It is possible to have a spontaneous shift from one energy state to another in quantum mechanics. For example, one can have an electron excited to a higher energy state by an incoming photon, and this is easy enough to understand in the context of mechanical event causality, since the change was sparked by an event. However, one can also have an electron in an excited state spontaneously emitting a photon and dropping back into the lowest energy state. There is no event that triggers this. It just happens. In terms of event causality, this makes no sense. Thus in terms of physics alone event causality is wrong (we have not ruled out a supra-physical or supernatural event causing the decay, though we haven't ruled it in either). But substance causality in physics remains: the efficient cause of the emitted electron and the low energy electron is just the high energy or excited electron. The laws of conservation of energy and momentum (which are still present in both quantum mechanics and quantum field theory) insist that such effects cannot appear out of nothing; there has to be some substance proceeding the effect. Furthermore in quantum mechanics energy and momentum are only transferred by quantum waves, whether the 'particles' such as electrons or 'force carrying particles' such as the photons that convey electromagnetic radiation. In other words, there is still a principle of causality describing this spontaneous emission, but the causality developed in response to the mechanical world-view is invalid; this does not mean that every type of causality is invalid.

10.3. Relativistic Quantum Mechanics

Non-relativistic quantum mechanics by itself is a highly successful theory, but it obviously suffers because it doesn't include special relativity: it breaks down when we approach the speed of light. So we need to construct a relativistic version of this theory (Aitchison and Hey, 2012; Strange, 2008; Bjorken and Drell, 1964). This might seem to be trivial: we just replace the Newtonian Energy equation with its special relativistic equivalent:

$$(\hat{H} - QA_0)^2 |\psi\rangle = (\hat{\vec{\mathbf{P}}} - Q\vec{\mathbf{A}})^2 c^2 |\psi\rangle + m^2 c^4 |\psi\rangle , \tag{325.1}$$

using the same operators as before and incorporating an electromagnetic potential $(A_0, \vec{\mathbf{A}})$ with electric charge Q. This is the Klein-Gordon equation, and it can be used (in fact it is used for spin 0 electrically neutral particles), though there are a few differences between this and the non-relativistic case, including a slightly different definition of the probability density; instead of $\psi^*\psi$ as in equation (312.2), it now needs to be $\psi^*(\hat{H} - QA_0)\psi$ if there is to be consistency between the change of the probability of a particle in a specific volume across time and the 'amount of probability' flowing in and out of that volume. Even though it can be used, it is harder to use than the non-relativistic Shrödinger equation. It is second order in the Energy, and non-linear, which means that most of the tricks used to solve the non-relativistic equation don't work for the Klein-Gordon equation. In particular, for every positive energy eigenstate of the Hamiltonian operator there is also a negative energy state. This is troubling, because it would suggest that the particles would tumble down to an energy of minus infinity. The new definition of probability density also causes problems when we have an electromagnetic field – it can be negative, which doesn't make any sense at all.

Dirac came up with another option. He noticed that if we replace the state $|\psi\rangle$ with a four component object called a spinor (mathematically similar to a vector, but given a different name to distinguish these four components from the indices of a vector which represent space/time directions), then it is possible to factorise the Klein Gordon equation into a simple linear equation, called the Dirac equation,

$$(\hat{H} - QA_0)\,|\psi\rangle = \vec{\alpha}\cdot(\hat{\vec{\mathbf{P}}}c - Q\vec{\mathbf{A}})\,|\psi\rangle + \beta mc^2\,|\psi\rangle\,, \tag{326.1}$$

where $\vec{\alpha}$ and β are 4×4 matrices ($\vec{\alpha}$ is a vector containing three 4×4 matrices) acting on the components of $|\psi\rangle$ which satisfy

$$\begin{aligned}\alpha_i\alpha_j + \alpha_j\alpha_i &= 2\delta_{ij}\\ \beta^2 &= 1\\ \alpha_i\beta + \beta\alpha_i &= 0.\end{aligned} \tag{326.2}$$

I implicitly multiply any scalars in the Dirac equation above by the 4×4 identity matrix, so $(\hat{H}-QA_0)$ really reads $(\hat{H}-QA_0)$ multiplied by the identity matrix, and the right hand side of the relationships between $\vec{\alpha}$ and β above should also be multiplied by the identity matrix. δ_{ij} is again a symbol which gives 1 if $i=j$ and 0 otherwise. The Dirac equation looks superficially similar to the Schrödinger equation, and it is based on similar principles. Once again we have the energy (or time evolution operator) acting on the likelihood of finding a particle at a location $(\vec{\mathbf{x}},t)$, $|\psi(\vec{\mathbf{x}},t)\rangle$ and the statement that this is equivalent to another set of operators including the electric potential, magnetic potential and momentum/translation operator acting on the same probability amplitude. One main difference is that we also have a term proportional to the particle's mass contributing to the energy: this is just the conclusion from special relativity that mass is a form of energy. The second difference from Schrödinger's equation is that this is a matrix equation. But nonetheless, it can be solved in a similar way to the Schrödinger equation, and one gets similar results, only better because one has now included special relativity.

It is possible to write the Dirac equation in a slightly different way, which more resembles the standard 4 vector notation used in special relativity. We write the differential four vector $\partial_\mu = (\frac{\partial}{\partial t}, c\vec{\nabla})$, and $\gamma^\mu = (\beta, -\beta\vec{\alpha})$ is not a four vector (because it doesn't transform in the same way from one inertial frame to another) but slots in the equations in the same way. This gives the form,

$$(\gamma^\mu(i\hbar\partial_\mu - QA_\mu) + m)\,|\psi\rangle = 0. \tag{326.3}$$

The Dirac equation avoids some of the problems of the Klein-Gordon equation. It has the same formula for the probability density as Schrödinger quantum mechanics, which means that we don't have to worry about negative probabilities or any nonsense such as that. However, it still admits to negative energy states. There were two ways around this. The first was to use the Pauli-exclusion principle, and note that you could only have one Fermion in each particle state. So if the all the negative energy states were filled up by matter we can't observe, then the ground state would be at energy zero as expected. The second is to propose the existence of the anti-particle. This interprets the negative energy states as positive energy anti-particles. These would be objects with the same mass as the original matter, but opposite electric charge. The spinor state $|\psi\rangle$ contains four components. However, the form of the γ matrices naturally splits these into two sets of two component states. One of these two component states appears to describe the particle (such as an electron), and the other the anti-particle (which should be called the anti-electron but is actually called the positron). The experimental discovery of the positron a few years after Dirac predicted its existence was another quantum mechanical triumph.

But it gets better. Spin interacts with magnetic fields, and to get this to agree with experiment it was necessary to incorporate spin into the Schrödinger equation in a fairly ad-hoc way. In fact, we need to write down two Schrödinger equations for a spin $\frac{1}{2}$ particle

such as the electron, one representing the spin $+\frac{\hbar}{2}$ state and the other the spin $-\frac{\hbar}{2}$ state (like angular momentum, the eigenstates of spin in quantum mechanics are separated by integer multiples of $\hbar$), and spin introduces a mixing between the two states. But in the Dirac equation, we have these two remaining components of the spinor describing the particle. Naturally, we can associate these with the spin states, and, as it happens, if you look at the non-relativistic limit of the Dirac equation (i.e. send the speed of light to infinity) you get precisely the non-relativistic Schrödinger equation including the additional fudge factor put in to account for spin. It's not just that we have to insert spin into the equations by hand to make them agree with experiment as in Schrödinger's quantum mechanics, in relativistic quantum mechanics you can't avoid getting it almost exactly[12] right. The Dirac equation describes a spin $\frac{1}{2}$ particle, and since most matter in the universe is spin $\frac{1}{2}$, the Dirac equation describes almost all matter in the universe. So allowing us to understand and account for spin was another major triumph of the Dirac equation. Once we understand spin, we also now fully understand magnets: a magnetic lump of iron has all the individual iron atoms in the same spin state; the tiny contribution from each atom combines, and the net result is a large and measurable magnetic field.

But there is a serious problem with the Dirac equation. It is to do with barriers of potential energy. The usual set-up in the text books is somewhat artificial, but the issue extends to more physical situations as well. We imagine that there is a sharp barrier of potential energy somewhere in space. In classical mechanics, this represents a sudden force encountered by a particle as it travels along. Something similar to a brick wall. You fire the projectile at the wall. If it is travelling slowly, then it bounces off the wall with (in the ideal system that doesn't correlate to real life, but is in all the textbooks because it's easier to compute) the same speed but in the opposite direction: it reflects off the wall. Alternatively, if the kinetic energy of the particle is larger than the energies binding the various bricks together, then it will smash through the wall, and continue in the same direction, albeit with a reduced velocity since some of its kinetic energy has gone into breaking up the wall. We call the case where it bounces off the wall reflection, and the second case either refraction or transmission depending on which physical system we are talking about. That's the classical picture. The quantum mechanical picture is, of course, different. In this case, some of the wavefunction $|\psi\rangle$ is transmitted and some of it is reflected. In other words, if you performed this experiment numerous times, sometimes you would see the particle bounce off the wall and sometimes you will see it transmit through the wall, with the relative frequencies of the transmitting and reflecting cases calculable from the probability distributions contained in the theory, which depend again on the particle's input kinetic energy (and in particular whether it is larger or smaller than the size of the potential barrier). In non-relativistic quantum mechanics, the probability for the reflection and the probability for the transmission add up to one. This is good. It means that the particle either transmits or reflects, and we don't have it suddenly vanishing or anything silly like that.

In relativistic quantum mechanics, we get a similar picture, at first. At incident energies smaller than the size of the potential wall, the particle reflects off the barrier (although the probability distribution extends a short distance into the barrier). Once the particle energy rises above the potential energy of the barrier, then we start seeing a signal for transmission, though part of the probability distribution reflects. The probability for transmission and reflection all add up to one; all in agreement with the non-relativistic case. Raise the energy further, and the reflection probability slowly decreases towards zero. But then, as we keep raising the energy of the input particle, something new happens. The probabilities for transmission and reflection no longer add up to one. They add up to three. We fire one particle into our experiment, and we have three coming out. The barrier itself hasn't changed: we are not blasting these particles out of the barrier; in this theoretical

[12]The 'almost' arises from small corrections from quantum field theory.

calculation the mathematical function representing the barrier just describes some idealised electromagnetic potential, without specifying what causes it. The barrier need not have within it any particles of the same type as the incoming particle. What is happening here? $E = mc^2$ is happening. The incoming particle kinetic energy which marks the transition between the one particle and three particle output states is twice the mass of the original particle. Kinetic energy is being converted into mass, with the potential acting as a catalyst. A new particle/anti-particle pair, travelling in opposite directions has come into existence. We have created new matter; not out of nothing (which is impossible) but with the kinetic energy of the incident particle being converted into the mass of the new matter. And, of course, this is not just something purely theoretical with no practical consequences; particle creation has been frequently observed experimentally.

This presents two problems. The first is philosophical. One of the premises of the mechanical world view was that corpuscles of matter couldn't be created or destroyed. That has to be thrown into the rubbish bin; the conversion of kinetic energy into new particles has been confirmed by many experiments. This was not really a concern for the twentieth century physicist, who was quite happy to just shut up and calculate and leave the philosophising to the philosophers. So we will leave that issue to one side. The second problem is a concern for the physicist. A quantum mechanical state $|\psi\rangle$ is meant to represent a single particle. Now we have three particles trying to squeeze into it. There is a mathematical inconsistency here, and mathematical consistency is one thing that all theoretical physicists do care about. This inconsistency is unavoidable in relativistic quantum mechanics.

The problem is not with the mathematics, but with quantum mechanics. It served us very well, it pointed us in the right direction, and it was very nearly there, providing us with much of what we need, but its not the right answer. We need something new with the creation and annihilation of particles right at its centre. Quantum mechanics needs to go the way of (the physicist part of) Aristotle, Newton and the dodo. It's time for something new: quantum field theory.

10.4. Quantum Field Theory

Of course, calling quantum field theory 'new' is a bit of an exaggeration. It's formalism and formulation is very much inspired by quantum mechanics (Peskin and Schroeder, 1995; Ryder, 1996; Weinberg, 1995; Bjorken and Drell, 1965). You need the same basic mindset to appreciate it. This encouraged most (from what I have seen) of the philosophers of twentieth century physics to start with quantum mechanics and then try to modify their ideas to incorporate field theory. To my mind, this is a mistake. Just as it is wrong to over-exaggerate the differences between quantum mechanics and quantum field theory, so it is wrong to suppose a continuity of ideas between the two. Field theory is its own thing; related to quantum mechanics as men are genetically, and in some respects behaviourally, similar to chimpanzees, but ultimately you are not going to fully understand human society from a detailed study of the chimpanzee.

In my view, the first of the important differences between the two theories as far as the philosophy is concerned concerns determinism. Quantum mechanics leaves us a little bit unsure about determinism; the wavefunction evolves deterministically, but observations lead to unpredictability. Quantum field theory completes this picture. We no longer have a deterministically evolving wavefunction. Physical determinism is an idea that needs to be rejected; it served us well by giving us Newton, but it's wrong and cannot be considered tenable in the light of modern physics. This, of course, also jettisons the event causality favoured by enlightenment philosophers (or those enlightenment philosophers who considered causality important). Quantum mechanics administered the poison to this idea; field theory upped the dose to fatal levels and paid for the funeral. However, causality isn't

thrown away entirely: Aristotelian substance causality still survives. It has to, because energy/momentum is still conserved. Every particle in a given state emerges from some other particle or particles.

The second difference that quantum mechanics retained from classical mechanics is the idea that matter is indestructible, and can be reduced to the sum of its parts. So let us consider the following scenario: we have an electron, happily going on its way, when all of a sudden it emits a photon. Where did that photon come from? The classical intuition can view this as though the photon was perhaps trapped inside the electron, as a worm might be inside an apple, then all of a sudden it comes to the surface and escapes. If the photon emerges, then it had to be part of the electron in the first place. This is derived from the notion that matter cannot be created or destroyed. But this picture is false. For one thing, the electron can, in principle, emit an infinite number of photons; one worm in an apple is believable, but an infinite number is not. Equally, the photon can subsequently decay into an electron/positron pair, or quark anti-quark pair (a completely unrelated set of particles), which makes little sense if it was originally hidden inside the electron in some way. Secondly, such a scenario can't explain circumstances where there is no electron left after the event, for example when an electron and positron annihilate to leave a single photon. Can we also say that hidden inside the photon that is hidden inside the electron are an actually existing infinite number of electron/positron pairs. In the beginning there is only the electron; there is no actual photon (whose presence is described by a very different mathematical operator).

Well what else could be happening? It is common among physicists to say that the photon, or the particle/anti-particle pair are pulled out of the vacuum. But that presents philosophical problems of its own. Isn't the vacuum meant to be nothing? Well, in reality, the vacuum of quantum field theory isn't nothing (the name 'vacuum' was applied before people properly understood what it was, and now we are stuck with the name even though the word is now used to describe something which is different from the meaning of the word that you will find in all the older dictionaries), but contains a great deal of structure in the underlying gluonic field. These structures exist, and they do affect observable physics (for example, they are to a large extent responsible for the confining force that holds quarks together into protons, neutrons and other particles). But even so, saying that these particles exist in the vacuum until they are pulled out of it by our first electron is troubling: particles carry energy (if nothing else their rest mass), and yet all the energy for the created particles comes from the kinetic energy of the original particle. The structures in the vacuum are limited to various objects composed of gluons. If there was also a background of the infinite number of particles required to explain the creation and annihilation of particles, then it would imply a universal background energy, which would be detected by gravitational experiments. It is computed that if this explanation were correct, the dark energy of the universe would be 120 orders of magnitude larger than is observed. Therefore the explanation is not correct; the created particle/anti-particle pair does not come from the vacuum.

So what then? We have no choice but to abandon the mechanistic principle that matter cannot be created or destroyed, or one type of matter cannot turn into another. The photon doesn't exist actually within the electron; we can, however, say that the electron in the lower energy state exists potentially within the initial electron and the photon virtually – in precisely the same sense as in Aristotelian metaphysics, as defined in chapter 4. The electron has the power to decay into a photon and another electron in a lower energy state. The photon has the power to decay into an electron positron-pair. Where we started with one particle, we now have three. The pair comes from the photon; in effect it turns into them - the only reason we are reluctant to admit this possibility is prejudice based on a lifetime of mechanistic indoctrination (not least because the mechanistic principle seems to apply so well at the sort of scales we are used to; its breakdown is in the microscopic world we cannot directly observe). Conservation of energy implies a certain degree of continuity between the initial and final states, and demands a principle of substance causality. And we

are now, of course, dangerously approaching an Aristotelian philosophy as the most natural interpretation of quantum field theory. But enough of this philosophising. Back to the physics.

Before proceeding, I need to make one important clarification. There are several different, but equivalent, ways of proceeding with quantum field theory, all of which make the same predictions at the end, but which at the start look completely different. In particular, I want to here mention the Hamiltonian and Lagrangian approaches to field theory, and in particular the different ways in which they treat time. The Hamiltonian picture essentially treats each moment of time as separate, and asks how the fields evolve in time. This is similar to how we think of things; we see the world as it is now, and ask what it will be like in the future. The past, present and future are seen as almost different worlds, though linked together by the time evolution operator, the Hamiltonian. We see the progress of time from one moment to the next, and imagine that that is what the universe is like. Many people think that only the present is real, with the past and future lost in the sea of potentialities. The problem with this picture is that it does not make the Lorentz invariance of special relativity – where time is treated in a very similar way to space manifest. It is conceptually neat, but mathematically a bit messy. The Lagrangian approach makes the Lorentz invariance obvious. One cannot get away from it; mathematically it is just there. I personally find this formulation of quantum field theory far more satisfying and natural. However, it is very different from our common-sense notions of time. The predictions of the two approaches are the same; indeed they are entirely equivalent once we translate everything from one 'language' to another. Here I will describe the Hamiltonian, or canonical, approach to quantum field theory, leaving the Lagrangian approach to a later chapter, since the Hamiltonian approach is a more obvious successor of (my discussion of) relativistic quantum mechanics.

My discussion of field theory will not be a formal and rigorous proof from the axioms below, simply because that would be too technical for my purposes. Such a proof is, of course, possible. Instead I will outline an argument which will give a flavour of the proof, and hopefully an understanding of how the theory arises from the premises and of the main features of the theory, while skipping over many of the details.

AXIOM 10.4.1.

(1) *I introduce the Fock state* $|n_1, n_2, n_3, n_4, \ldots\rangle$ *where each* n_i *refers to the number of particles in an energy state* i*. So the state given in the previous sentence means that there are* n_1 *particles in state* 1, n_2 *particles in state* 2, n_3 *particles in state* 3 *and so on (having previously determined and numbered the possible physical states of the system). As in quantum mechanics, this represents a complex*[13] *probability amplitude describing our knowledge of the system.*

(2) *I assume, as in quantum mechanics, that observables are measured by operators acting on Fock states, and the measured values correspond to eigenstates of these operators. The notation used for the states will be as in quantum mechanics, with the same properties of completeness and orthonormality.*

(3) *I assume that the eigenstates of the Hamiltonian operator,* $\hat{H}$*, represent the energy states, and* $-\frac{i}{\hbar}\hat{H}$ *is also the time evolution operator for the Fock state.*

(4) *I assume that the momentum states are represented by the differential operator* $i\hbar\vec{\nabla}$*.*

(5) *I assume that there is an operator for each type of particle which describes the creation of that particle with momentum* $\vec{\mathbf{p}}$*, which I denote as* $\hat{a}^\dagger_{\vec{\mathbf{p}}}$ *multiplied by a constant of proportionality, and an operator* $\hat{a}_{\vec{\mathbf{p}}}$ *which describes the annihilation of a particle with the same momentum (again, baring a constant of proportionality*

[13] *Complex* means the mathematical definition that it can have both real and imaginary parts.

which I will determine later)[14] *Other types of particles might require creation and annihilation operators* $\hat{b}^\dagger_{\vec{\mathbf{p}}}$ *and* $\hat{b}_{\vec{\mathbf{p}}}$*, and so on. The creation and annihilation operators are Hermitian conjugates of each other. Suppose that the creation operator acting on the state* $|0\rangle$ *gives the state* $|1\rangle$*:*

$$\langle 1|\,\hat{a}^\dagger\,|0\rangle = \lambda_+, \tag{331.1}$$

for some normalisation λ_+*. If we take the Hermitian conjugate of this equation, we get,*

$$\langle 0|\,\hat{a}\,|1\rangle = \lambda_+^*, \tag{331.2}$$

which just states that the operator $\hat{a}$ *acting on the state* $|1\rangle$ *gives something proportional to the state with no particles. But that is precisely what the annihilation operator does. Thus if* $\hat{a}^\dagger$ *is the creation operator, then* $\hat{a}$ *is proportional to the corresponding annihilation operator. Since the normalisation of these operators (the constants* λ_+*) is something we will derive later, we can without loss of any generalisation define* $\hat{a}$ *to be our annihilation operator, and let any constant of proportionality be sorted out later (according to the definition in the next axiom).*

(6) The number operator is given by $\hat{a}^\dagger_{\vec{\mathbf{p}}}\hat{a}_{\vec{\mathbf{p}}}$ *for particles of the type described by the creation operator* $\hat{a}$ *and* $-\hat{a}^\dagger_{\vec{\mathbf{p}}}\hat{a}_{\vec{\mathbf{p}}}$ *for the corresponding anti-particles.*

(7) I assume that the system is indeterminate, i.e. particles can be created or destroyed unpredictably but only in various proscribed ways limited by (for example) various conservation laws. Thus there is a certain likelihood which can be known for each type of decay or recombination of particles in a given process and time interval. Thus given an initial state probability amplitude, we can compute a final state probability amplitude for various different possible final states using the time evolution operator.

(8) We assume that there are three space dimensions and one time dimension with a hyperbolic geometry, with different inertial frames connected by Lorentz symmetries (i.e. special relativity).

(9) I assume local gauge invariance, so $e^{i\alpha(x)}\,|n_1, n_2, n_3, n_4, \ldots\rangle$ *is physically equivalent to* $|n_1, n_2, n_3, n_4, \ldots\rangle$ *(since these lead to the same probabilities).*

(10) Each type of particle has a particular mass, which satisfies (in the absence of the potential) $E^2 - p^2c^2 = m^2c^4$*. The particles also have a definite electric charge.*

(11) For spin $\frac{1}{2}$ *fields, the energy operator is as defined by the Dirac equation; for spin 0 particles one needs the Klein-Gordon equation.*

(12) The fundamental material particles, electrons, quarks and so on, are spin $\frac{1}{2}$ *particles with the experimentally observed masses and electric charges.*

(13) There are also gauge fields with vector potentials (and polarisations) similar to that described in classical electrodynamics (though with their own creation/annihilation operators).

(14) I assume that the low momentum theory is independent of any very high momentum physics.

---

[14] I use the terms *creation* and *annihilation* because they are the standard terminology in the particle physics literature. However, like many names chosen before the thing they refer to was fully understood, they might be misleading. In popular parlance, the terms creation and annihilation often represent the emergence of particles from *nothing* or their destruction into *nothing*. However, as we shall see, in this case that idea is mistaken. The physical process involves one type of substance transforming into another. The philosophical names for this type of substantial change are *generation* and *corruption*, and it would be more consistent to name these operators as *generation* and *corruption* operators. The terminology used in physics thus contradicts the terminology used in philosophy for the same process. As a physicist, I will follow the language of my own field, and when I use the terms *creation* and *annihilation* in this context, they should be treated as synonyms for the philosophical *generation* and *corruption*.

(15) I assume that the interaction of particles is point-like: there is no interaction at a distance.

(16) I assume local isolation of particles, i.e. we can study individual particles and their interactions without reference to what is happening in the rest of the universe. The interaction between two or more particles can be described wholly by referencing the initial and final states of those particles during that interaction, and does not depend on the states of any other matter in the universe.

These axioms are either drawn directly from experiment, or are made in response to experimental results; some just define various representational conventions[15]. Their justification arises in part from how well the theory built from these premises conforms with experiment, and in part that many of the axioms, such as the spin of the particles, gauge fields, and the number of space-time dimensions are directly observed. The notation is carried on from quantum mechanics, with many of the same ideas, but is adapted to a theory where the particles are not indestructible but can be extracted from or converted into kinetic energy. The idea is that there is a real world system out there, and our knowledge of it should be distinguished from the system. The system is indeterminate as the creation and destruction of particles is spontaneous, so our knowledge of the system cannot in principle be perfect, as after we have made a measurement there is a certain likelihood that it would have changed through spontaneous emission and we have no means of predicting how the system will evolve. Measurements involve comparison between the state of the universe and a trial state, and connect our knowledge with the real world. They update our knowledge of the system, and thus modify the amplitude that describes the knowledge.

The Fock state as defined is a definite eigenstate of the number operator, which counts how many particles in each state. However, we can instead express it as counting the number of particles in the eigenstates of any operator. The use of the number operator is entirely arbitrary. We conventionally normalise the Fock state to 1, so

$$\langle n_1, n_2, n_3, n_4, \ldots | n_1, n_2, n_3, n_4, \ldots \rangle = 1. \tag{332.1}$$

Equally, the states are orthogonal, so

$$\langle n_1', n_2', \ldots | n_1, n_2, \ldots \rangle = 0 \tag{332.2}$$

if any of the number of particles represented by n' are not equal to the corresponding number of particles represented by n. In particular, I will introduce the vacuum state $|0\rangle = |0, 0, 0, \ldots\rangle$, i.e. the state with no particles in it. The creation operator acting on this vacuum state gives a state with one particle, so, for example, $\hat{a}_1^\dagger |0\rangle = \lambda_+ |1, 0, 0, \ldots\rangle$ and so on, for some coefficient λ_+ which needs to be determined but can't be zero. This is reversed by the annihilation operator $\hat{a} |1, 0, 0, \ldots\rangle = \lambda_- |0\rangle$ for another coefficient λ_-. Any annihilation operator acting on the vacuum state gives 0, $\hat{a}_{\vec{\mathbf{p}}} |0\rangle = 0$. These states basically are the quantum field theory equivalent of the quantum mechanical wavefunction, only that they describe the whole system rather than just individual particles.

CONJECTURE 10.4.2. *The creation and annihilation operators do not commute, so* $\hat{a}_{\vec{\mathbf{p}}}^\dagger \hat{a}_{\vec{\mathbf{p}}} \neq \hat{a}_{\vec{\mathbf{p}}} \hat{a}_{\vec{\mathbf{p}}}^\dagger$

Proof:
The annihilation operator acting on the vacuum state gives zero, and thus $\hat{a}_{\vec{\mathbf{p}}}^\dagger \hat{a}_{\vec{\mathbf{p}}} |0\rangle = 0$. However, the annihilation operator acting on the one particle state does not give zero, thus

[15] For example, it can be shown that $\hat{a}\hat{a}^\dagger$ should be proportional to the number operator; saying that it is the number operator merely fixes the normalisation of $\hat{a}$ and $\hat{a}^\dagger$. If we put in a constant of proportionality, then other equations would similarly be adjusted so that the experimental predictions are the same. We just have to choose some number, and it ultimately makes no difference what we choose. So why not make life easier for ourselves by choosing the simplest option?

$\hat{a}_{\vec{p}}\hat{a}^\dagger_{\vec{p}}|0\rangle = \lambda_+\lambda_-|0\rangle$. Thus $(\hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}} - \hat{a}_{\vec{p}}\hat{a}^\dagger_{\vec{p}})|0\rangle \neq 0$, which shows that the operators do not commute. ■

As stated in the discussion of location and momentum wavefunctions in the section on quantum mechanics, we can rotate the creation and annihilation operators to a new basis,

$$\begin{aligned}\hat{a'}^\dagger_i &= U_{ij}\hat{a}^\dagger_j \\ \hat{a'}_i &= U^\dagger_{ij}\hat{a}_j\end{aligned} \tag{333.1}$$

for some unitary (i.e the Hermitian conjugate $U^\dagger$ of U is its inverse; this is the complex analogue of a rotation) operator U. To keep the notation simpler, I have used discrete indices i, j in place of the momentum, but the same principle will apply for a continuous momentum variable.

CONJECTURE 10.4.3. *For $i \neq j$, $\hat{a}^\dagger_i\hat{a}^\dagger_j = c_{ij}\hat{a}^\dagger_j\hat{a}^\dagger_i$ where the coefficient $c_{ij} = \pm 1$.*

Proof:
The first element of this proof is to remark that the two states must be proportional to each other. Acting on some input state $|n_1, n_2, n_3, \ldots, n_i, \ldots, n_j, \ldots\rangle$, $a^\dagger_j$ increases n_j by 1 and $a^\dagger_i$ increases $n^\dagger_i$ by one. Thus

$$\begin{aligned}\hat{a}^\dagger_i\hat{a}^\dagger_j|n_1, n_2, n_3, \ldots, n_i, \ldots, n_j, \ldots\rangle &\propto |n_1, n_2, n_3, \ldots, n_i + 1, \ldots, n_j + 1, \ldots\rangle \\ \hat{a}^\dagger_j\hat{a}^\dagger_i|n_1, n_2, n_3, \ldots, n_i, \ldots, n_j, \ldots\rangle &\propto |n_1, n_2, n_3, \ldots, n_i + 1, \ldots, n_j + 1, \ldots\rangle\end{aligned} \tag{333.2}$$

and thus the two sets of operators must be proportional to each other. This means that

$$\hat{a}^\dagger_i\hat{a}^\dagger_j|0\rangle = c_{ij}\hat{a}^\dagger_j\hat{a}^\dagger_i|0\rangle, \tag{333.3}$$

for some coefficient c_{ij}; the ij subscript shows that this value might depend on which two particles we are creating. Writing down this equation with the i and j indices reversed means that $c_{ji} = \frac{1}{c_{ij}}$. The only question is what this constant of proportionality is. If we rotate to another basis, then we find that

$$\hat{a'}^\dagger_k\hat{a'}^\dagger_l U^\dagger_{ki}U^\dagger_{lj}|0\rangle = c_{ij}\hat{a'}^\dagger_k\hat{a'}^\dagger_l U^\dagger_{kj}U^\dagger_{li}|0\rangle \tag{333.4}$$

and, in particular,

$$(\hat{a}^\dagger_i\hat{a}^\dagger_j - c_{ij}\hat{a}^\dagger_j\hat{a}^\dagger_i)|0\rangle = 0 = U^\dagger_{ki}U^\dagger_{lj}(\hat{a'}^\dagger_k\hat{a'}^\dagger_l - c_{ij}\hat{a'}^\dagger_l\hat{a'}^\dagger_k)|0\rangle, \tag{333.5}$$

so that $c_{ij}c_{lk} = 1$. But our choice of index for l and k is arbitrary, so we could write this equation with them the other way round, so $c_{ij}c_{kl} = 1$. But this just means that $c_{kl} = c_{lk}$. However, c_{kl} is also $1/c_{lk}$, which is only possible if $c_{kl} = \pm 1$. Thus the conjecture is proved. ■

We can also write down a similar proof in the case that $i = j$. In this proof, I used the vacuum state $|0\rangle$. Does that restrict the applicability of the result? No, because we can use any Fock state, and it doesn't change the argument. We need to put something there, and I used $|0\rangle$ as an example, but the result is completely general.

We define the notation $[\hat{A}, \hat{B}]$ to mean $\hat{A}\hat{B} - \hat{B}\hat{A}$ and $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$. In this case, we have two options for creation operators. The first we call Boson fields, where

$$[\hat{b}^\dagger_i, \hat{b}^\dagger_j] = 0. \tag{333.6}$$

The second are called Fermion fields, where

$$\{\hat{a}^\dagger_i, \hat{a}^\dagger_j\} = 0. \tag{333.7}$$

This implies that $\hat{a}^\dagger_i\hat{a}^\dagger_i = 0$, i.e. you can't have two Fermions in the same state, which gives us the quantum mechanical Pauli exclusion principle. It also means that for Fermions each

n_i can only be 0 or 1; we need $a_i^\dagger |n_i = 1\rangle = 0$. By a similar process of reasoning, we find that for the annihilation operators,

$$\begin{aligned}[\hat{b}_i, \hat{b}_j] =& 0 \\ \{\hat{a}_i, \hat{a}_j\} =& 0.\end{aligned} \tag{334.1}$$

We now just have to find the commutation or anti-commutation relationships between the creation and annihilation operators.

CONJECTURE 10.4.1. *For Bosons, creation and annihilation operators have the commutation relation,* $[\hat{b}_i, \hat{b}_j^\dagger] \propto \delta_{ij}$. *For Fermions, creation and annihilation operators have the anti-commutation relation* $\{\hat{a}_i, \hat{a}_j^\dagger\} \propto \delta_{ij}$.

Proof:
The number operator was defined as $\hat{b}_i^\dagger \hat{b}_i$, so $\langle\psi| \hat{b}_i^\dagger \hat{b}_i |\psi\rangle$ counts how many particles (Bosons) of type b in state i are in the Fock state $|\psi\rangle$. Suppose that $|\psi\rangle$ contains no particles of type b_i; in this case the number operator correctly gives 0. The state $|n_i = 1\rangle$ has one particle in the state i, and this is proportional to $b_i^\dagger |0\rangle$. This means that the number operator acting on $\hat{b}_i^\dagger |0\rangle$ must give $\hat{b}_i^\dagger |0\rangle$, i.e this is an eigenstate of the number operator with eigenvalue 1. $\hat{b}_i^\dagger \hat{b}_i^\dagger |0\rangle$ must be an eigenstate of the number operator with eigenvalue 2, and so on. Now if we suppose the commutation relation, $[\hat{b}_i, \hat{b}_i^\dagger] = 1$ then, because $b_i |0\rangle = 0$, this works in every case,

$$\begin{aligned}\hat{b}_i^\dagger \hat{b}_i \hat{b}_i^\dagger |0\rangle =& \hat{b}_i^\dagger [\hat{b}_i, \hat{b}_i^\dagger] |0\rangle + \hat{b}_i^\dagger \hat{b}_i^\dagger b_i |0\rangle = \hat{b}_i^\dagger |0\rangle \\ \hat{b}_i^\dagger \hat{b}_i \hat{b}_i^\dagger \hat{b}_i^\dagger |0\rangle =& \hat{b}_i^\dagger [\hat{b}_i, \hat{b}_i^\dagger] \hat{b}_i^\dagger |0\rangle + \hat{b}_i^\dagger \hat{b}_i^\dagger [\hat{b}_i, \hat{b}_i^\dagger] |0\rangle + \hat{b}_i^\dagger \hat{b}_i^\dagger \hat{b}_i^\dagger \hat{b}_i |0\rangle = 2\hat{b}_i^\dagger \hat{b}_i^\dagger |0\rangle ,\end{aligned} \tag{334.2}$$

and so on. Only this commutation relation gives the right result for each state of the number of particles we need. $\hat{b}_j^\dagger \hat{b}_i^\dagger |0\rangle$ should also be an eigenstate of the number operator (as indeed with any combination of b_j operators), which requires that the number operator for state i must commute with the creation operator for state j. This implies $[\hat{b}_j, \hat{b}_i^\dagger] = 0$. Thus for Bosons, we find,

$$[\hat{b}_i, \hat{b}_j^\dagger] = \delta_{ij}. \tag{334.3}$$

There is a similar argument for Fermions. Since the number operator needs to commute with $\hat{a}_j^\dagger$ and $\hat{a}_i^\dagger$ anti-commutes with $\hat{a}_j^\dagger$, we need $\hat{a}_i$ to also anti-commute with $\hat{a}_j^\dagger$. We thus might expect to find for Fermions,

$$\{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij}. \tag{334.4}$$

In practice, the number operator is a little bit different when we have particles and anti-particles present. The number we need to measure is the number of particles minus the number of anti-particles, $\hat{a}_j^\dagger \beta \hat{a}_j$, where β is the matrix that gives -1 if there is an anti-particle and 1 if there is a particle. It acts on the (suppressed) spinor indices in the $\hat{a}$ and $\hat{a}^\dagger$ operators. This is the same β matrix that appeared in the Dirac equation. This slightly complicates matters, since we want to keep the notation as similar as possible for all types of field. One way of doing this is to write the creation operator as $\hat{a}^\dagger \beta$ instead of $\hat{a}^\dagger$. We are free to do this: this still acts on the vacuum state to give a one particle state, which is all we need for the creation operator. If we do this, then the number operator goes back to being the creation operator times the annihilation operator, which means that we can continue to use our earlier arguments. ■

I started this section with a discussion of discrete states, however in reality we are interested in states distinguished by a continuous momentum index $\vec{\mathbf{p}}$. The reason I started with a presentation of the theory for discrete states is that similar concepts apply in the

continuous case, and the proofs follow the same principles, only the argument is a little more technical, and thus one tends to become bogged down in the technical details instead of the underlying principles, and it is the underlying principles I am most interested in. However, the continuous and discrete systems are not completely analogous, so I will spend a moment now stating the modifications we need to the above commutation relations to apply them to a continuous set of states.

The extension to continuous momenta rather than discrete indices uses Dirac δ-functions, so the natural way to extend the discussion above is to write

$$\{\hat{a}_{\vec{\mathbf{p}}}, \hat{a}^\dagger_{\vec{\mathbf{p}}'}\beta\} = \delta^3(\vec{\mathbf{p}} - \vec{\mathbf{p}}'). \tag{335.1}$$

However, there is a problem here. This particular construction keeps time and space separate, and also momentum and energy. However, we want the final result to be in agreement with special relativity. That means we need to take extra care in some places to make sure that we satisfy the Lorentz symmetries that underlie special relativity. This is one of those places. The equation above is not Lorentz invariant. However, it can be shown (I'm not going to do so here, but the proof is reasonably easy) that $E_{\vec{\mathbf{p}}}\delta^3(\vec{\mathbf{p}} - \vec{\mathbf{p}}')$ is Lorentz invariant, where $E_{\vec{\mathbf{p}}}$ (or p_0) is the energy of the state created by $a^\dagger_{\vec{\mathbf{p}}}$. This means that we really need the expression to be

$$\{\hat{a}_{\vec{\mathbf{p}}}, \hat{a}^\dagger_{\vec{\mathbf{p}}'}\beta\} = p_0\delta^3(\vec{\mathbf{p}} - \vec{\mathbf{p}}'). \tag{335.2}$$

Alternatively, we can see that $\frac{1}{\sqrt{p_0}}\hat{a}_{\vec{\mathbf{p}}}$ and $\frac{1}{\sqrt{p_0}}\hat{a}^\dagger_{\vec{\mathbf{p}}}$ satisfy the 'standard' anti-commutation relation. So we can also use $\frac{1}{\sqrt{p_0}}\hat{a}_{\vec{\mathbf{p}}}$ as our annihilation operator and $\beta\frac{1}{\sqrt{p_0}}\hat{a}^\dagger_{\vec{\mathbf{p}}}$ as our creation operator, and keep the easier anti-commutation rule. This also applies to the various Boson fields: we need to use $\frac{1}{\sqrt{p_0}}\hat{b}^\dagger_{\vec{\mathbf{p}}}$ as the creation operator for a b-particle, and so on. This fixes the constant of proportionality for the creation and annihilation operators.

These creation and annihilation operators are for states of definite momentum. It is useful to also have a creation operator, or field, representing a state of definite location. As in quantum mechanics, the momentum operator is the differential with respect to location, so we can expect the location creation operator to be something along the lines of $e^{-i\vec{\mathbf{p}}\cdot\vec{\mathbf{x}}}\hat{a}^\dagger_{\vec{\mathbf{p}}}$.

$$\hat{\psi}^\dagger_s(\vec{\mathbf{x}}) = \int d^3p \frac{1}{(2\pi)^3\sqrt{p_0}}\hat{a}^\dagger_{\vec{\mathbf{p}},s'}\beta_{s',s}e^{-i\vec{\mathbf{p}}\cdot\vec{\mathbf{x}}}. \tag{335.3}$$

Spin $\frac{1}{2}$ particles are described by four component spinors, and the s represents the index telling us which component if the four component spinor it is. Two components of the spinor field refer to a particle, and the other two to the anti-particle. For the Boson field (which is its own anti-particle), it is conventional to write the location operator in a symmetric way.

$$\hat{\phi}(\vec{\mathbf{x}}) = \int \frac{d^3p}{(2\pi)^3\sqrt{p_0}}(\hat{b}^\dagger e^{-i\vec{\mathbf{p}}\cdot\vec{\mathbf{x}}} + \hat{b}e^{i\vec{\mathbf{p}}\cdot\vec{\mathbf{x}}}) \tag{335.4}$$

There should also be creation and annihilation operators for the photon field representing the electromagnetic potential. In this case, because of the gauge symmetry, there is not a single unique way in which this can be done. One common option is to use the gauge freedom to set A_0 to zero. This leaves us with,

$$\hat{\vec{\mathbf{A}}}(x,t) = \int \frac{d^3k}{(2\pi)^3\sqrt{2\hbar\omega}}\left(\hat{\mathbf{A}}_{\vec{\mathbf{k}}\alpha}\xi^\alpha e^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{x}}-i\omega t} + \hat{\vec{\mathbf{A}}}^\dagger_{\vec{\mathbf{k}}\alpha}\xi^{\alpha\dagger}e^{-i\vec{\mathbf{k}}\cdot\vec{\mathbf{x}}+i\omega t}\right) \tag{335.5}$$

for an electromagnetic wave with frequency ω and polarisation ξ. Here $\frac{1}{\sqrt{\hbar\omega}}\hat{\vec{\mathbf{A}}}$ is the annihilation operator, and $\frac{1}{\sqrt{\hbar\omega}}\hat{\vec{\mathbf{A}}}^\dagger$ the creation operator.

CONJECTURE 10.4.5. *For a Boson field, acting on a state $|n_i = m\rangle$ with m particles in state i, $\hat{b}^\dagger_i|n_i = m\rangle = \sqrt{m+1}\,|n_i = m+1\rangle$, where $|n_i = m+1\rangle$ is a state with $m+1$ particles in*

state i.

Proof:
We know that $\hat{b}_i^\dagger |n_i = m\rangle$ is proportional to the state $|n_i = m+1\rangle$. The question is what is the constant of proportionality. So let us write, $\hat{b}_i^\dagger |n_i = m\rangle = \alpha |n_i = m+1\rangle$. Taking the Hermitian conjugate of this expression gives $\langle n_i = m| \hat{b}_i = \langle n_i = m+1| \alpha^*$. Putting these together gives

$$\langle n_i = m| \hat{b}_i \hat{b}_i^\dagger |n_i = m\rangle = \alpha^* \alpha \langle n_i = m+1 | n_i = m+1\rangle. \tag{336.1}$$

With the state $\langle n_i = m+1 | n_i = m+1\rangle$ normalised to 1, and $\hat{b}_i \hat{b}_i^\dagger = 1 + \hat{b}_i^\dagger \hat{b}_i$, with $\hat{b}_i^\dagger \hat{b}_i$ the number operator so $\langle n_i = m| \hat{b}_i^\dagger \hat{b}_i |n_i = m\rangle = m$, we have $\alpha^* \alpha = m+1$. The complex phase of this quantity is arbitrary, so we might as well make it a real number. Thus $\hat{b}_i^\dagger |n_i = m\rangle = \sqrt{m+1}\, |n_i = m+1\rangle$. ■

The calculation for the Fermion field is similar, giving $\hat{a}_i |n_i = 1\rangle = |n_i = 0\rangle$.
So, having got the players in place, we are now ready to turn to the actual dynamics.

CONJECTURE 10.4.6. *An operator with eigenvalues $E_{\vec{\mathbf{p}}}$ acting on a Fermion field which commutes with the number operator can be written as*

$$\hat{O} = \int \frac{d^3p}{(2\pi)^3 p_0} \hat{a}_{\vec{\mathbf{p}}}^\dagger \beta E_{\vec{\mathbf{p}}} \hat{a}_{\vec{\mathbf{p}}}. \tag{336.2}$$

Proof:
We can see this by considering what happens when we apply this operator to a Fock state $|\psi\rangle$. The expected value of the operator is $\langle \psi | \hat{O} | \psi \rangle$. This should give the sum of the values of $E_{\vec{\mathbf{p}}}$ for each particle present in the state, and nothing for those particles which are absent. The representation of $\hat{O}$ above clearly does this. If the particle is absent, then $\hat{a}_{\vec{\mathbf{p}}} |\psi\rangle$ gives 0. On the other hand, if it is present, then $\hat{a}_{\vec{\mathbf{p}}} |\psi\rangle$ will give the state $|\psi'\rangle$ which is the state ψ with that particle missing. $E_{\vec{\mathbf{p}}}$ then records the value of the eigenvalue, and then $\hat{a}_{\vec{\mathbf{p}}}^\dagger |\psi'\rangle E_{\vec{\mathbf{p}}} = |\psi\rangle E_{\vec{\mathbf{p}}}$. Applying $\langle \psi|$ to this gives $E_{\vec{\mathbf{p}}}$, and we sum up over all possible states. Thus

$$\langle \psi | \int \frac{d^3p}{(2\pi^3 \sqrt{p_0}} \hat{a}_{\vec{\mathbf{p}}}^\dagger E_{\vec{\mathbf{p}}} \hat{a}_{\vec{\mathbf{p}}} |\psi\rangle = \langle \psi | \sum_{\text{particles in } \psi} E_{\vec{\mathbf{p}}} |\psi\rangle, \tag{336.3}$$

which is just what we expect from $\langle \psi | \hat{O} | \psi \rangle$. This is true for any possible state $|\psi\rangle$, which proves the conjecture. ■

This is obviously true for the energy operator, $\hat{H}$, since the total energy of the system is just the sum of the energies of each particle in the system.

CONJECTURE 10.4.7. *There exists a time evolution operator $e^{\frac{i}{\hbar}\Delta t \hat{H}}$ which describes how the state evolves from a time t to a time $t + \Delta t$, where Δt is some infinitesimal time interval. This time evolution operator can be written as $\hat{H} = \int \frac{d^3p}{(2\pi)^3} \frac{\beta \hat{a}_{\vec{\mathbf{p}}}^\dagger}{\sqrt{E_{\vec{\mathbf{p}}}}} E_{\vec{\mathbf{p}}} \frac{\hat{a}_{\vec{\mathbf{p}}}}{\sqrt{E_{\vec{\mathbf{p}}}}}$ for some $E_{\vec{\mathbf{p}}}$ where $\vec{\mathbf{p}}$ labels various eigenstates of the time evolution operator (i.e. meta-stable states).*

Proof:
We observe that change occurs in nature; therefore the Fock state should be written as a function of time, for example as $|\psi(t)\rangle$. The state at a time $t + \Delta t$ is therefore $|\psi(t + \Delta t)\rangle$. Expanding this in a Taylor series gives $|\psi(t + \Delta t)\rangle = e^{\Delta t \frac{\partial}{\partial t}} |\psi(t)\rangle$. We can define the

operator $\hat{H} = -i\hbar\frac{\partial}{\partial t}$, and thus

$$|\psi(t+\Delta t)\rangle = e^{\frac{i}{\hbar}\Delta t\hat{H}}|\psi(t)\rangle . \tag{337.1}$$

We can expand $|\psi(t)\rangle$ in the eigenstates of the operator $\frac{\partial}{\partial t}$, and we label these eigenstates with the variable $\vec{\mathbf{p}}$, and the corresponding eigenvalues are defined to be $E_{\vec{\mathbf{p}}} = p_0$. These eigenstates represent particles which are meta-stable in time, except for (unimportant physically) changes in the phase of the corresponding Fock state. We define $\beta\frac{\hat{a}^\dagger_{\vec{\mathbf{p}}}}{\sqrt{E_{\vec{\mathbf{p}}}}}$ as the appropriate creation operator for these states. Using the formalism developed above, we can therefore write

$$\begin{aligned}|\psi(t_0+\Delta t)\rangle =& T[e^{i\Delta t\hat{H}/\hbar}]\,|\psi(t_0)\rangle \\ =& T[e^{i\Delta t\int\frac{d^3p}{2\pi p_0}\hat{a}^\dagger_{\vec{\mathbf{p}}}\beta E_{\vec{\mathbf{p}}}\hat{a}_{\vec{\mathbf{p}}}}]\,|\psi(t_0)\rangle ,\end{aligned} \tag{337.2}$$

where T is the same time-ordering operator introduced in the discussion of quantum mechanics (equation (312.3)). ■

Discussion:
Since the energy operator is defined as the generator of time translations, this means that the energy eigenstates (the states unchanged under application of the energy operator) describe the stable states of the system, which represent the physical substances. The energy operator shares the same eigenstates as the momentum operator, but the location operator has different eigenstates. This means that the stable states of the system, and consequently the physical substances, should be parameterised in terms of the momentum rather than the location. This is the reason why I am concentrating on operators, $\hat{a}^\dagger_{\vec{\mathbf{p}}}$, creating states of definite momentum. While it is certainly possible to define use an operator $\hat{a}^\dagger_{\vec{\mathbf{x}}}$ to represent the creation of a particle with definite location, it is better (especially in a summary such as this) to concentrate on the physically relevant states. □

CONJECTURE 10.4.8. *Lorentz invariance and gauge invariance mean that we can also express $E_{\vec{\mathbf{p}}}$ as a function of the eigenvectors of the spatial derivative operator $i\hbar\vec{\nabla}$, which I denote as $\hat{\vec{\mathbf{p}}}$. For a spin $\frac{1}{2}$ field (described by a four component spinor), this allows us to write the time evolution operator in the form*

$$E_{\vec{\mathbf{p}}} = \vec{\alpha}\cdot(\hat{\vec{\mathbf{p}}} - Q\hat{\vec{\mathbf{A}}}) + \beta m + Q\hat{A}_0, \tag{337.3}$$

where $(\hat{A}_0, \hat{\vec{\mathbf{A}}})$ are combined annihilation/creation operators for the field representing the electromagnetic four-potential, m the mass of the particle, and Q the electric charge of the particle.

Proof:
This form of the energy is forced on us by the requirements of Lorentz invariance (i.e. to satisfy special relativity and respect the hyperbolic geometry of space/time) and gauge invariance (i.e. the phase of the Fock state is not physical). That it represents the only way to represent the time derivative operator in terms of spatial derivatives (with $\vec{\mathbf{p}}$ the eigenvector of $i\hbar\vec{\nabla}$) was shown above in the discussion of the Dirac equation. We still need to show that gauge invariance requires the presence of the electromagnetic operator $\hat{A}_\mu$ as in equation (337.3).

If the Fock state including this phase is written as $e^{i\theta}|\psi\rangle$, then the differential of $e^{i\theta}|\psi\rangle$ is

$$\partial_\mu(e^{i\theta}|\psi\rangle) = e^{i\theta}\left(e^{-i\theta}\partial_\mu(e^{i\theta})|\psi\rangle + \partial_\mu|\psi\rangle\right). \tag{337.4}$$

We define $iQ\hat{A}_\mu$ in such a way so that when the phase θ is changed then $iQ\hat{A}_\mu \to iQ\hat{A}_\mu - \frac{1}{\hbar}e^{-i\theta}\partial_\mu(e^{i\theta})$.

Ultimately, the phase θ doesn't contribute to any expectation value; this means that we want it to cancel out of quantities such as $\langle\psi|\,\hat{O}\,|\psi\rangle$ for any operator $\hat{O}$. We should get the same result whether or not we multiply the Fock state by the phase $e^{i\theta}$, which is just an artefact of our choice of abstraction and not physical. An operator whose expectation value is not a function of the gauge is known as gauge covariant. Only gauge covariant operators represent physical observables, that is quantities that have meaning when we map back from the mathematical abstraction to reality. However, as shown above, the operator ∂_μ is not gauge covariant, and thus cannot be used in any computation. But $i\hbar\partial_\mu - Q\hat{A}_\mu$ is gauge covariant, and therefore this can represent a physical observable. We therefore need to incorporate $\hat{A}_\mu$ into every derivative operator. The twin requirements of gauge covariance and Lorentz symmetry therefore restrict us to write the eigenvalues of the stationary states in the form (337.3) for a spin $\frac{1}{2}$ particle. This is actually the same form as the Dirac equation.

This leaves open the question of what to do with the operator $\hat{A}_\mu$ which we have had to introduce to preserve local gauge invariance. This is an operator which acts on the Fock state, which means that it can be constructed from creation and annihilation operators (since any local operator – which concerns individual particles rather than the system as a whole – either creates a particle, destroys a particle, or is proportional to the number operator in some basis) – but not the creation and annihilation operators for the Fermions, because since it follows the Fermion annihilation operator there are no Fermions around for it to act on. Therefore it represents the creation or annihilation operator of some different type of particle. In practice, this turns out to be a spin 1 Boson field, which we call the photon. As with the Fermions, to preserve Lorentz invariance the creation and annihilation operators need to be normalised by the square root of their energy. I'll label the momenta of the photons as k_μ.

The difference between this analysis and relativistic quantum mechanics is that in quantum mechanics $\vec{\mathbf{A}}$ is treated as a constant background mathematical function, while in field theory $\hat{\vec{\mathbf{A}}}$ is treated as the creation operator for a different type of particle.

Therefore the time evolution operator for the Fermion part of a Fock state can be written as

$$|\psi(t_0+\Delta t)\rangle = T[e^{i\Delta t\int \frac{d^3p}{(2\pi)^3 p_0}\hat{a}^\dagger_{\vec{\mathbf{p}}}\beta(\vec{\alpha}\cdot(\hat{\vec{\mathbf{p}}}-Q\frac{\hat{\vec{\mathbf{A}}}}{\sqrt{k_0}})+\beta m+Q\frac{\hat{A}_0}{\sqrt{k_0}})\hat{a}_{\vec{\mathbf{p}}}}]\,|\psi(t_0)\rangle\,. \tag{338.1}$$

We also need to consider the time evolution of the photon field. Once again, this is constrained by Lorentz and gauge invariance. We start by constructing a quantity that contains the creation and annihilation operators, and which satisfies Lorentz and gauge invariance, and this is $(F_{\mu\nu}[\hat{A}])^2 = (\partial_\mu\frac{\hat{A}_\nu}{\sqrt{k_0}} - \partial_\nu\frac{\hat{A}_\mu}{\sqrt{k_0}})(\partial^\mu\frac{\hat{A}^\mu}{\sqrt{k_0}} - \partial^\mu\frac{\hat{A}^\nu}{\sqrt{k_0}})$, which we set to be some constant value 2Λ. This is known as the Maxwell action, since it is built from the Lagrangian that generates Maxwell's equations. Acting on momentum states, this gives

$$\frac{\hat{A}_{\vec{\mathbf{k}},\mu}}{\sqrt{k_0}}(\eta^{\mu\nu}k^2 - k^\mu k^\nu)\frac{\hat{A}_{\vec{\mathbf{k}},\nu}}{\sqrt{k_0}} = \Lambda \tag{338.2}$$

so the energy k_0 sandwiched between the creation operators for photons with that energy is

$$\hat{H}_g = \frac{\hat{A}_{\vec{\mathbf{k}},\mu}}{\sqrt{k_0}}k_0\eta^{\mu\nu}\frac{\hat{A}_{\vec{\mathbf{k}},\nu}}{\sqrt{k_0}} = \Lambda + \frac{\hat{A}_{\vec{\mathbf{k}},\mu}}{\sqrt{k_0}}(\frac{k^\mu k^\nu + \eta^{\mu\nu}\vec{\mathbf{k}}^2}{k_0})\frac{\hat{A}_{\vec{\mathbf{k}},\nu}}{\sqrt{k_0}} \tag{338.3}$$

because $k^2 \equiv k_\mu k^\mu = k_0^2 - \vec{\mathbf{k}}^2$.

We can add this into our time evolution operator, giving the final result

$$|\psi(t_0+\Delta t)\rangle = T[e^{i\Delta t\hat{H}}]\,|\psi(t_0)\rangle \tag{338.4}$$

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3p d^3q}{(2\pi)^6\sqrt{p_0 q_0}} \hat{a}^\dagger_{\vec{\mathbf{p}}} \beta \Big(\vec{\alpha} \cdot (\vec{\mathbf{p}}\delta^3(\vec{\mathbf{p}} - \vec{\mathbf{q}})\delta^3(\vec{\mathbf{k}}) - Q\delta^3(\vec{\mathbf{p}} + \vec{\mathbf{k}} - \vec{\mathbf{q}}) \frac{\hat{\vec{\mathbf{A}}}_{\vec{\mathbf{k}}}}{\sqrt{k_0}}) + \delta^3(\vec{\mathbf{p}} - \vec{\mathbf{q}})\delta^3(\vec{\mathbf{k}})\beta m + Q \frac{\hat{\vec{\mathbf{A}}}_{\vec{\mathbf{k}}0}}{\sqrt{k_0}} \delta^3(\vec{\mathbf{p}} + \vec{\mathbf{k}} - \vec{\mathbf{q}}) \Big) \hat{a}_{\vec{\mathbf{q}}} + \int \frac{d^3k}{(2\pi)^3 k_0^2} \hat{\vec{\mathbf{A}}}_{\vec{\mathbf{k}},\mu} (k^\mu k^\nu + \eta^{\mu\nu}\vec{\mathbf{k}}^2) \hat{\vec{\mathbf{A}}}_{\vec{\mathbf{k}},\nu} + \Lambda. \tag{339.1}$$

This is a bit of a handful, but follows directly from the presentation so far. $\hat{H} = \hat{H}_f + \hat{H}_g$ represents two terms; the first describes how the probability amplitude for the Fermion fields evolve in time, and the second how the probability amplitude for the photon fields evolves in time. I have inserted Dirac δ-functions (as defined below equation (214.3)) to demonstrate the conservation of momentum when the photon and Fermion fields interact; these arise from some of the details which I have skipped over in this presentation. Otherwise this is just taken from equations (338.1) and (338.3). It is heavily constrained by Lorentz and gauge invariance. This is the simplest expression we can write down consistent with these two axiomatic principles. It is not the only expression we can write down consistent with these two principles – this will become important when we discuss re-normalisation a little later - but the other possible expressions are related to it through rotations of the basis of the various creation and annihilation operators (as in equation (333.1)). I will proceed by first of all using this expression and seeing where it leads us, and then discussing what happens if we rotate the basis for the creation and annihilation operators. ■

CONJECTURE 10.4.9. *The probability that the state at the later time is in a state $\langle\psi_1|$ is the modulus square of the amplitude*

$$P(\psi_1 \setminus \psi(t_0), \Delta t) = \frac{\langle\psi(t_0)| T[e^{-i\Delta t\hat{H}/\hbar}] |\psi_1\rangle \langle\psi_1| T[e^{i\Delta t\hat{H}}] |\psi(t_0)\rangle}{\sum \psi_i \langle\psi(t_0)| T[e^{-i\Delta t\hat{H}/\hbar}] |\psi_i\rangle \langle\psi_i| T[e^{i\Delta t\hat{H}}] |\psi(t_0)\rangle}, \tag{339.2}$$

where $\sum \psi_i$ indicates the sum over every possible state.

Proof:
This was derived above in the discussion of quantum mechanics (around equation 312.2). ■

We can write the state $\psi(t_0)$ as the product of creation operators acting on the vacuum state $|0\rangle$.

So suppose that we have a initial set of particles with momenta $\vec{\mathbf{p}}_1$, $\vec{\mathbf{p}}_2$ and so on, and we want to find the probability that we finish with a set of particles with momenta $\vec{\mathbf{k}}_1$, $\vec{\mathbf{k}}_2$, and so on. This is then dependent on

$$\frac{\langle 0| T \left[\frac{\hat{a}_{\vec{\mathbf{k}}_1}}{\sqrt{E_{\vec{\mathbf{k}}_1}}} \frac{\hat{a}_{\vec{\mathbf{k}}_2}}{\sqrt{E_{\vec{\mathbf{k}}_1}}} \dots e^{i\frac{1}{\hbar}\int dt \frac{d^3p}{(2\pi)^3} \frac{d^3k}{(2\pi)^3} \hat{I}_{\vec{\mathbf{p}},\vec{\mathbf{k}}}} \frac{\hat{a}^\dagger_{\vec{\mathbf{p}}_1}}{\sqrt{E_{\vec{\mathbf{p}}_1}}} \beta \frac{\hat{a}^\dagger_{\vec{\mathbf{p}}_2}}{\sqrt{E_{\vec{\mathbf{p}}_2}}} \beta \dots \right] |0\rangle}{\langle 0| T \left[e^{i\frac{1}{\hbar}\int dt \frac{d^3p}{(2\pi)^3} d^3k \hat{H}_{\vec{\mathbf{p}},\vec{\mathbf{k}}} + F_{\mu\nu}[\hat{A}]^2} \right] |0\rangle}. \tag{339.3}$$

Why use the vacuum state $|0\rangle$ here? This is the state of the universe with no particles in it, so, once we apply our creation operators to introduce the initial states of the particles we are studying, we seem to be saying that there are only these two (or three, or however many we study) particles in the universe, an obvious absurdity. This is where the assumption of local separation comes in, i.e that the rest of the universe does not affect our experiment. In practice, of course, we should introduce creation operators for every other particle in the universe – or sum over a weighted distribution of those creation operators since we don't

know the wave-functions of these particles but these will only interfere with our system through electromagnetic interactions, and with numerous competing effects from differently charged particles we can assume that this background noise will be sufficiently small that their effects will be well below our experimental error. So using the vacuum state here is acceptable.

Equation (339.3) is related to the probability: the probability is the modulus square of this probability amplitude. The denominator provides the required normalisation to ensure that the total probability for all possible physical processes adds up to one. The Hamiltonian operator links Fermion operators describing particles with two different momenta, the difference between the momenta of the two Fermions being the momenta of the gauge field $\hat{\vec{\mathbf{A}}}$:

$$\hat{H}_{\vec{\mathbf{p}},\vec{\mathbf{k}}} = \frac{1}{\sqrt{E_{\vec{\mathbf{p}}}}}\hat{a}^{\dagger}_{\vec{\mathbf{p}}}\beta\left[(i\vec{\alpha}\cdot\vec{\mathbf{p}}+\beta m)\delta(\vec{\mathbf{p}}-\vec{\mathbf{k}}) - \frac{1}{\sqrt{E^{A}_{\vec{\mathbf{p}}-\vec{\mathbf{k}}}}}Q\vec{\alpha}\cdot(\hat{\vec{\mathbf{A}}}_{\vec{\mathbf{p}}-\vec{\mathbf{k}},\alpha}\xi^{\alpha} + \hat{\vec{\mathbf{A}}}^{\dagger}_{\vec{\mathbf{k}}-\vec{\mathbf{p}},\alpha}\xi^{\alpha})\right]\frac{1}{\sqrt{E_{\vec{\mathbf{k}}}}}\hat{a}_{\vec{\mathbf{k}}}. \tag{340.1}$$

There are three terms which contribute to this Fermion energy operator. The first is the mass (βm), which is straight-forward. The second is the momentum term in the energy ($\vec{\alpha}\cdot\vec{\mathbf{p}}$), coming from the derivative operator in the Dirac Hamiltonian. The third term features the gauge field, which, as with all other quantum fields, needs to be expressed in terms of creation and annihilation operators. When we convert from the location space representation of the Hamiltonian operator to the momentum space representation, we pick up a factor of $\int d^4x e^{ix^{\mu}\sum p_{\mu}}$, where the sum is over the momentum of every particle present (positive for the created particles, negative for the destroyed particles), which are the two creation and annihilation operators and the gauge fields. The expression arises in this form because the Dirac operator is local. There is a mathematical relationship which converts this to a delta function over all the momenta $\int d^4x e^{ix^{\mu}\sum p_{\mu}} = (2\pi)^4\delta(\sum p_{\mu})$, which means that only the solution with $\sum p_{\mu} = 0$ contributes. This equation says that the sum of all the momenta for the created particles minus the destroyed particles is zero; it demands the conservation of energy and momentum. Inside the $\hat{H}$ operator, this means that the momentum for the photon has to be the difference between the created and destroyed Fermions, so if the photon field is created, then its momentum plus the momentum associated with the $\hat{a}^{\dagger}$ creation operator must equal the momentum associated with the $\hat{a}$ annihilation operator. Similarly, if the photon is destroyed, then the destroyed Fermion plus destroyed photon must have the same energy/momentum as the created Fermion. This explains the form of the operator $\hat{H}$.

CONJECTURE 10.4.10. *The probability amplitude describes a system in which numerous particles are created and annihilated. The probability amplitude for a physical process can be written as the sum of the amplitudes for each way in which one can create and destroy particles such that, starting from the stated initial state, one reaches the given final state.*

Proof:
We can expand the exponential in equation (339.3) describing the time evolution of the probability amplitude in a Taylor series. This will give us is a long string of creation and annihilation operators with various momenta and various integrals over the momenta. For example, we can consider a case where the initial state (which I will call A) contains two Fermions, and the final state (which I will call B) contains two Fermions. So we have two particles to start with, and we finish with two particles. You might think that this is straight-forward: if we were discussing mechanics this would just represent the collision between two objects, and with conservation of momentum the problem can be solved easily. Not so in quantum field theory: this simple process can involve the creation and destruction of a huge number of particles (anything up to infinity) before the final state is reached.

The numerator for the first terms in the expansion for two Fermions to two Fermions with different momenta reads (to just give one of the leading terms),

$$\langle 0|\, T\Bigg[\frac{a_{\vec{k}_1}}{\sqrt{E_{\vec{k}_1}}}\frac{\hat{a}_{\vec{k}_2}}{\sqrt{E_{\vec{k}_2}}}$$
$$\left(1-\frac{i}{\hbar}\left(\int\frac{d^3q_1}{(2\pi)^3}\int d^3q_2 Q\frac{\hat{a}^\dagger_{\vec{q}_1}}{\sqrt{E_{\vec{q}_1}}}\beta\vec{\alpha}\cdot\left(\frac{\hat{\vec{\mathbf{A}}}_{\vec{q}_2-\vec{q}_1,\alpha}\xi^\alpha}{\sqrt{E^A_{\vec{q}_2-\vec{q}_1}}}+\frac{\hat{\vec{\mathbf{A}}}^\dagger_{\vec{q}_1-\vec{q}_2,\alpha}\xi^\alpha}{\sqrt{E^A_{\vec{q}_1-\vec{q}_2}}}\right)\frac{\hat{a}_{\vec{q}_2}}{\sqrt{E_{\vec{q}_2}}}\right)\right.$$
$$\left.-\frac{1}{2\hbar^2}\left(\int\frac{d^3q_1}{(2\pi)^3}\int d^3q_2 Q\frac{\hat{a}^\dagger_{\vec{q}_1}}{\sqrt{E_{\vec{q}_1}}}\beta\vec{\alpha}\cdot\left(\frac{\hat{\vec{\mathbf{A}}}_{\vec{q}_2-\vec{q}_1,\alpha}\xi^\alpha}{\sqrt{E^A_{\vec{q}_2-\vec{q}_1}}}+\frac{\hat{\vec{\mathbf{A}}}^\dagger_{\vec{q}_1-\vec{q}_2,\alpha}\xi^\alpha}{\sqrt{E^A_{\vec{q}_1-\vec{q}_2}}}\right)\frac{\hat{a}_{\vec{q}_2}}{\sqrt{E_{\vec{q}_2}}}\right)^2+\ldots\right)$$
$$\frac{a^\dagger_{\vec{p}_1}}{\sqrt{E_{\vec{p}_1}}}\beta\frac{a^\dagger_{\vec{p}_2}}{\sqrt{E_{\vec{p}_2}}}\beta\Bigg]|0\rangle\,. \quad (341.1)$$

The full expansion looks horrible, and that's because it is. There are an infinite number of terms, all a mass of creation and annihilation operators. But there are some simplifying features and patterns. The $\frac{a^\dagger_{\vec{p}_1}}{\sqrt{E_{\vec{p}_1}}}\beta\frac{a^\dagger_{\vec{p}_2}}{\sqrt{E_{\vec{p}_2}}}\beta\,|0\rangle$ on the right of the equation represents our initial state: we have two creation operators acting on the vacuum which represent the two initial particles. On the left of the equation, we have $\langle 0|\,\frac{a_{\vec{k}_1}}{\sqrt{E_{\vec{k}_1}}}\frac{\hat{a}_{\vec{k}_2}}{\sqrt{E_{\vec{k}_2}}}$. These two annihilation operators represent our probe to measure the possible final states of the system. The part in the middle represents the dynamics, and here we have numerous terms containing expressions such as $\hat{a}^\dagger(\hat{\vec{\mathbf{A}}}^\dagger+\hat{\vec{\mathbf{A}}})\hat{a}$. One interpretation of this is that we can annihilate one Fermion, and create another Fermion and either create or destroy a photon. In other words, this represents a process where a Fermion either emits or absorbs a quanta of electromagnetic radiation. Alternatively, recall that the Fermion annihilation operator also can represent the creation of an anti-Fermion. This would represent a photon splitting into two particles: a Fermion and anti-Fermion (this is the process which allows kinetic energy to be converted into mass: a photon is emitted from a Fermion, carrying away energy and momentum, and then decays into a pair of particles). Equally, we can have a Fermion and anti-Fermion annihilate into a photon.

The way the creation and annihilation operator algebra works is that every creation operator has to be paired up with an annihilation operator with the same momentum. We start with a vacuum state, and then create a bunch of particles and destroy a bunch of particles, and then compare the final result of this process of creation and annihilation against the vacuum state. If we create a particle we don't destroy (or try to destroy a particle we don't create), then the final result will be not be a vacuum, so when we apply the $\langle 0|$ comparison operator we will just get zero, which won't contribute to the amplitude. For example, $\langle 0|\, a_{\vec{\mathbf{p}}}a^\dagger_{\vec{\mathbf{q}}}\,|0\rangle = 0$ unless $\vec{\mathbf{p}}=\vec{\mathbf{q}}$; one needs an annihilation operator to cancel out each creation operator. So for every particle we create with a certain momentum, we must also have the same particle destroyed somewhere in the expansion (which includes the creation and annihilation operators representing the initial and final states). So what we wind up with is an expansion where particles are constantly created and destroyed according to the four processes of photon emission, photon absorption, photon decay into particle and anti-particle pairs, and a particle and anti-particle pair destroying each other and emitting a photon. Each of these processes contribute a fixed factor to the final amplitude. They occur spontaneously, but with fixed likelihoods. We cannot tell which of these processes occur, but we can calculate the probability amplitude for each of them. Summing up these amplitudes allows us to compute the probability that we will finish with a given final state. In all of these processes, conservation of energy, momentum and electric charge is guaranteed.

The second part of the process involves the propagation of a particle with a given momentum. The equation above expanded the interaction part of the time evolution operator; we also have the non-interacting part, which for Fermions resembles $e^{i\frac{1}{E_{\vec{q}}}\hat{a}^\dagger_q\beta(i\vec{\alpha}\cdot\vec{q}+\beta m)\hat{a}_q}$. Expanding this gives terms which represent the continual spontaneous destruction and instantaneous re-creation then destruction and instantaneous re-creation, and so on, of a single particle (with the same momentum each time). This might seem like a bit of a waste, but each time we create and destroy the particle we pick up a factor of the particle momentum divided by the energy. So once again, the more times we go through the destruction/creation process, the less likely it is that the process occurs. So it is possible that the destruction/creation process occurs no times, once, twice, thrice, four times, and so on, and we have to add up contributions from each of these possibilities into our amplitude. Fortunately, this time (unlike the interacting term) the series takes a known mathematical form (a geometrical series[16]), and we can perform the sum to infinity without having to make any approximations. This then gives the contribution to the amplitude for the particle to travel directly with a certain momentum; the result is called the propagator.

What about all the numerous particle creation and and annihilation operators in the expansion that are completely isolated from our initial and final states? So, for example, we might have a annihilation operator from the expansion of the exponential absorbing an creation operator also from the exponential, with no interaction with our (physical) initial or final states. This would represent a particle appearing out of 'nothing' (in reality, it is not nothing, but the quantum electrodynamics vacuum, which is certainly something) and then disappearing again without interacting with anything else. As it happens, though, the sum of the terms involving apparent vacuum creation/destruction contributes multiplicity to the numerator and also appears in the denominator of the amplitude, which means that the two contributions cancel out completely. Vacuum creation does not contribute to any calculation of an observable in quantum field theory. There is thus a question about whether such vacuum creation of particle/anti-particle pairs is a real physical effect or just an accidental by-product of the mathematical formalism. Certainly it does not have any physical consequences, at least this side of a unification with gravity. Therefore, there is no way we can say for certain whether such spontaneous vacuum creation/annihilation happens in practice; we lose nothing by assuming that it doesn't. My own view is that this possibility is just an artefact of the mathematics, and particles are only created from the decay of other existent particles, or destroyed by absorption or decay into other existent particles. The physically significant quantity is the probability amplitude, not its numerator or denominator treated in isolation, and such vacuum creation/annihilation does not contribute to the amplitude. Thus it has no physical significance, and it is most natural to say that such processes are not present in reality.

Thus quantum field theory (or in this case quantum electrodynamics, since so far I have only discussed the theory of electromagnetism and matter) is simply a theory that treats mathematically the spontaneous emission, absorption and decay of particles. It gives us a set of rules describing which changes are possible, which decays are impossible, and a means to calculate the probability of the change. Each term in equation (341.1) represents one possible route in which these processes can combine to get from A to B. We do not

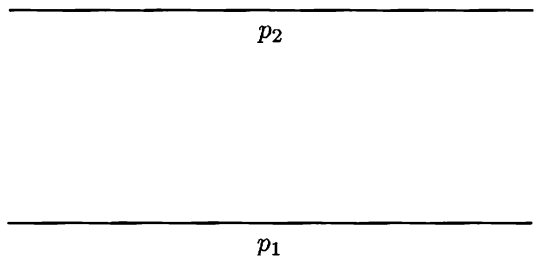
[16] A geometrical series is a series where each term is higher power of some coefficient, so, for example, $1 + a + a^2 + a^3 + a^4 + \ldots$. The expansion of the exponential gives $e^{c^\dagger_{\vec{p}}\hat{C}c_{\vec{p}}} = 1 + c^\dagger_{\vec{p}}\hat{C}c_{\vec{p}} + (c^\dagger_{\vec{p}}\hat{C}c_{\vec{p}})^2/2! + (c^\dagger_{\vec{p}}\hat{C}c_{\vec{p}})^3/3! + \ldots + (c^\dagger_{\vec{p}}\hat{C}c_{\vec{p}})^n/n! + \ldots$, where I am using $c^\dagger_{\vec{p}}$ to represent a generalised creation operator and C as the function of the momentum that sits between the creation and annihilation operators in the Hamiltonian. $n!$ is the factorial function, $n! = 1\times2\times3\times\ldots\times(n-1)\times n$, which represents the number of ways one can arrange n indistinguishable objects. When performing the operator algebra, we need to pair up each creation operator with a annihilation operator, but it doesn't matter which annihilation operator we choose. In the term containing n creation and annihilation operators, we therefore have $n!$ ways in which we can perform this pairing. This factor of $n!$ neatly cancels out the $1/n!$ contribution in the expansion of the exponential, leaving us with just a geometrical series in $\hat{C}$.

know, and cannot even in principle predict, the microscopic details of which particular process occurred in practice: that uncertainty is in part why we can only consider quantum field theory probabilistically. But nonetheless, because we can treat it probabilistically, the process is mathematically tractable.

So which of these 'routes' do we have to consider when computing the probability of a particular final state occurring given the initial state and these rules? The answer is all of them. Any of these infinite number of routes from A to B is possible. However, they are not all equally likely, which means that we can obtain increasingly accurate approximations to the final answer, as described below. ■

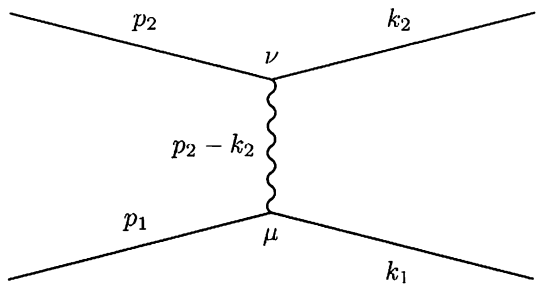
CONJECTURE 10.4.11. *It is possible to describe the possible processes of creation and annihilation diagrammatically.*

Discussion:
The diagrammatic representation of quantum field theory was developed by the great Richard Feynman, one of the pioneers of field theory. Each possible route from A to B is represented by its own diagram. The diagrammatic representation is completely equivalent to the algebraic representation above, but makes it easier to visualise the process of creation and annihilation of particles. The rules of the diagrams mean that a straight line represents a Fermion field, while a wiggly line represents a photon field. So, for example, the following diagram represents the two Fermions travelling straight across without interacting at all:



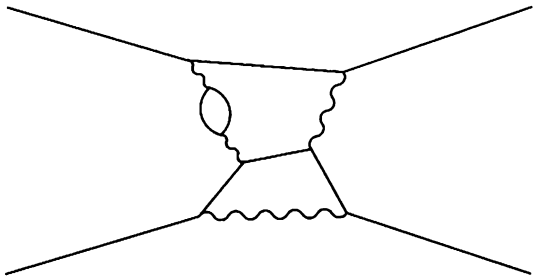
In this diagram, the direction from left to right indicates the direction of efficient causality. We start on the left with the causes (or initial states), we finish on the right with the effects (or final states). The labels represent the momenta of the two particles (it is possible to draw diagrams based around location rather than momentum, but in practice when solving them one converts to the momentum diagrams. Since the momentum states are the metastable states of the system, and location states aren't, it is not only easier but also better to focus on the momentum diagrams). The question we wish to answer is: what is the probability that we will end up with that final state given the initial state and the various axioms behind quantum field theory? In this particular diagram, not a lot happens: the two particles continue unhindered, and we finish in the same state that we started.

But, of course, this is not the only diagram we can draw to get from A to B. This next diagram represents the simplest form of interaction, where one of the Fermions emits a photon, which is subsequently absorbed by the other Fermion:



This diagram corresponds to the quadratic term in equation (341.1). Note that momentum is conserved, so $p_2 - k_2 = k_1 - p_1$: the photon carries momentum away from one Fermion and gives it to the other one.

But, of course, there are a lot more possibilities. Here is another one (I have left off the labels this time):

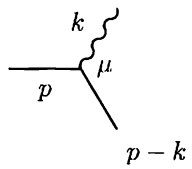


And we could go on drawing diagrams until we get bored of them. (Indeed, I did get bored of them when I was a student.) We can draw a photon line between any two Fermion lines, and that leads to a large number of possibilities.

Of course, the diagrams themselves are not much use unless we can make predictions from them; but fortunately we can do that. Each point and line on the diagram has a precise mathematical meaning. Each point where the photon line interacts with the Fermion lines is called a vertex, and each line is called a Fermion or photon propagator. We haven't yet discussed the photon propagator, but it arises from the Maxwell action in a similar way that the Fermion propagator arises from the Fermion momentum and mass term. The only complication is the gauge symmetry: as discussed before, the vector potential was undefined up to a derivative term: A_μ and $A_\mu + \partial_\mu \alpha$ for some function α of location are physically equivalent. It doesn't matter which gauge we use (i.e. which function α), but we have to be consistent throughout the calculation, and it is thus customary to ensure consistency by fixing the gauge using some additional arbitrary condition for A_μ that in effect selects the choice of α. In the propagator listed below, this freedom in choosing the gauge symmetry is described by expressing the propagator in terms of some constant called ζ. We select one particular ζ when we perform the calculation, and that selects the gauge (i.e. the complex phase in front of the Fock state). The final full expansion will not depend on ζ, but the intermediate steps of the calculation will so we have to keep track of it. We start a Fermion line for every particle in our initial state, and end a Fermion line for every particle in our final state. We label each line with the corresponding momentum, and each vertex with a direction. We then apply the following rules:

$$\underset{k}{\longrightarrow} = \frac{i}{\gamma^\mu k_\mu - im + i\epsilon} \qquad \text{(Fermion line)}$$

$$\mu \underset{k}{\sim\sim\sim} \nu = \frac{i}{k^2 + i\epsilon}(\eta_{\mu\nu} - (1-\zeta)\frac{k_\mu k_\nu}{k^2}) \qquad \text{(photon line)}$$



$= iQ\gamma^\mu$ (photon-Fermion vertex).

ϵ is just an infinitesimally small number; it tells us how to correctly handle any situations where the denominator of propagator becomes zero. A denominator becoming zero is usually bad in mathematics because it leads to an infinity, but in this case because it is within an integral it doesn't lead to any singularity in the final result; the $i\epsilon$ is just a bookkeeping measure to keep the intermediate steps of the calculation finite and ensure that the singularities are correctly treated in the calculation. At the end of the calculation, we send ϵ to zero. This way of regulating the symmetry actually, by the wonders of complex analysis, ensures that the final result is correctly time-ordered. We draw every possible diagram consistent with our desired input and output state, enforce conservation of momentum at each vertex, and integrate over every undetermined momentum[17].

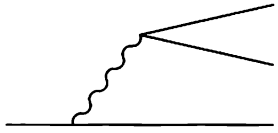
Each Feynman diagram represents one possible physical process, and by applying these Feynman rules we can compute the likelihood for that process. To obtain the overall likelihood for a particular final state, we simply sum up the contributions from each process that results in that final state. We cannot know which process occurred in practice, but we can compute likelihoods for each process, and from those likelihoods obtain probabilities. Note however that the energy and momentum of the internal lines are not necessarily bound to the particle's mass. In quantum field theory mass has to be considered simply as a parameter describing the self interaction of particles in a given momentum state, and it loses its direct meaning from special relativity (including relativistic quantum mechanics). The equation $E^2 = p^2c^2 + m^2c^4$, and the corresponding interpretation of the mass, only arises in the classical $\hbar \to 0$ limit of quantum field theory.

To compute the probability amplitude for a diagram, we need to assign the appropriate numerical factor to each line and vertex. Sometimes some of the momenta are undetermined, in which case we integrate over each momenta - this corresponds to summing over every possible way in which the process can happen. To calculate the complete amplitude we need to sum over every possible process which gives us the desired output state from the input state. This means summing over every possible diagram and every possible internal momenta, and, unfortunately, there are an infinite number of these which makes the full calculation both tedious and practically impossible. However, each interaction or vertex between Fermions and photons occurs with a certain probability amplitude, each propagation occurs with a probability amplitude which depends on the momentum, and within a single route these contributions to the probability amplitude multiply together. In particular, each vertex is multiplied by a factor which is (when you put in all the correct factors of $\hbar, \epsilon_0$ and c) much smaller than 1, so the more vertices a diagram has, the less it contributes to the overall result. This means that the more decays a 'route' has, the less likely it is to occur. We can thus class the different possibilities according to their likelihood, calculating the contributions to the amplitude from the most likely processes first, and then the ones in the next tier down, and so on. The easiest diagrams have no vertices, and next we have the diagram with two vertices shown above (and possibly one other two vertex diagram, depending on the types of Fermions involved). So if we just compute these diagrams, we will get an answer for the probability amplitude which is actually reasonably good. We can improve this answer by calculating the diagrams with 4 vertices, and improve it still further

[17]There is also an additional 'symmetry factor' which multiplies some diagrams, but since I'm not going to actually do the calculation, and this makes no difference to my overall argument, I will not discuss this further. I have also for the sake of simplicity neglected the polarisation of the photon fields in this discussion, which alters the equations a little bit, but doesn't change the overall picture: one simply has to add polarisation vectors to each photon field.

by computing the diagrams with 6 vertices, and so on. In each case, we know roughly how big the next terms in the series are going to be, how many diagrams there are and approximately what each diagram will contribute, so we can estimate the order of magnitude of the error which arises from truncating the series. Thus while we cannot compute the probability that an initial state goes to a permitted final state with perfect precision, we can nonetheless obtain an exceptionally good precision – all we need is that the theoretical uncertainty is smaller than the experimental uncertainty – and (just as importantly) compute a good estimate of the magnitude of the remaining error. If we need to be more precise, 'all' we have to do is compute the next terms in the series.

Note also that particle number is not necessarily conserved in these processes. An example is the diagram below, which represents the creation of a particle/anti-particle pair (as seen in the Klein paradox discussed above).



□

This is very nearly the theory of quantum electrodynamics, but it is not quite right. There is one additional problem. I noted before that mathematically we were free to rotate the basis of the creation and annihilation operators. So given one set of operators $\hat{a}^\dagger_{\vec{\mathbf{p}}}$, we can construct a second set of creation operators $\hat{a'}^\dagger_{\vec{\mathbf{p}'}} = \int d^3pU(\vec{\mathbf{p}'}, \vec{\mathbf{p}})\hat{a}^\dagger_{\vec{\mathbf{p}}}$, and the mathematical formalism restricts the choices of the rotation U a little but still leaves us with a great deal of freedom. In the calculation so far, I have chosen one basis where the energy operator is particularly simple and thus convenient. But does the creation operator $\frac{1}{\sqrt{p_0}}\hat{a}^\dagger_{\vec{\mathbf{p}}}$ actually create a physical electron, or some other particle? Could it be that one of the other possible $\hat{a'}^\dagger_{\vec{\mathbf{p}'}}$ operators corresponds to the actual electron we observe in our experiments instead? At first glance, there is no obvious way in which we can tell which operator is correct. However, there is one important feature that the correct theory will have, and this restricts our choices. That feature is that the computation gives finite probabilities. I mentioned that we have to integrate over all possible momenta for those momenta which are undetermined by the conservation of momenta. In principle, we need to integrate all the way up to infinite momenta, but these integrals don't converge if we do so: they give infinity. So what we might want to do[18] is to artificially introduce a maximum allowed momentum (called the momentum cut-off) for the particles, and then rotate the basis for our creation operators so that no single amplitude computed from one of the infinite number of physical processes we have to consider depends on this cut-off. We can then send the momentum cut-off to infinity safely. This leaves us with one single solution[19] for the basis that gives a finite result. This particular rotation of the basis (and corresponding rotation of the mass m

[18]This isn't actually what is done in practice, because a momentum cut-off violates the gauge symmetry, so in real life we have to be more sophisticated than this and use a means of regulating the theory that is consistent with both gauge and Lorentz symmetry. The same basic principles apply as in my discussion here, but the details are more complicated. Since my purpose is to explain the principle rather than go into detail, I will describe the simplest approach.

[19]Again, I am being a little too simplistic here. In practice, we have to make several arbitrary choices during renormalisation, such as the regularisation scheme. Thus, when we present a result, we present it alongside which re-normalisation scheme was used to calculate it. The situation is similar to the case in special relativity, where we have to state our coordinate system as well as the actual coordinates. We know how to convert from one scheme system to another, and the scheme is needed both in theoretical computation and interpreting the experiment. So as long as we consistently use the same scheme, there is no ambiguity in the comparison between theory and experiment. This freedom of re-normalisation scheme again relates to choices we make when mapping from the real world to a mathematical abstraction; it is not part of the real world, and is removed when we map back from the abstraction so we can compare with the

and the gauge field operator) corresponds to the physical electron. This process is called re-normalisation, and allows us to make computations. However, re-normalisation doesn't affect the underlying philosophical foundation of quantum field theory; it is basically a tool used to match up the creation and annihilation operators in our abstract representation of physics with the actual physical particles we observe.

The initial and final states in the expansion must correspond to the physical particle. This forces us to rotate the creation operators in the time evolution operator, so that the same basis is used throughout the calculation.

The perturbation expansion was performed with a free, non-interacting, Fermion as its basis. But in reality, there is no such object. The Fermion cannot be separated from the electromagnetic field: it continually emits and absorbs photons, and the carriers of the Weak force (and, more indirectly other matter fields and the gluons that carry the strong nuclear force, which can enter through higher order loop interactions). Effectively the particle we observe in nature is the sum of contributions of the bare electron, the bare electron emitting and absorbing a photon, and ultimately an infinite number of other terms. Not surprisingly, if we try to use a particle which doesn't and can't physically exist as the basis of our calculation, we will get a nonsensical result. The infinities are a result of us performing computations for an unphysical particle. However, there is some physical state which we observe as the electron in nature, and there will be some creation operator for this state, and if we used this state in our calculations it will give a mathematically consistent result. Because the creation and annihilation operators form a complete basis, we know that the creation operator for the physical particle will be some linear superposition of the bare electron operators. We can find that particular combination by the condition that probabilities should be finite, and (baring some ultimately unimportant convention-dependence) this solution turns out to be unique for each type of particle. Applying this transformation of basis at the end of the calculation is relatively straight-forward, and can be done order by order in the expansion used to perform calculations. Thus the final result is unique, well defined, mathematically consistent, and can be compared against experiment.

Note that a particle's interactions cannot be separated from the definition of the particle. The interactions affect the way in which it has to be renormalized, and thus the definition of the physical particle necessarily contains information on the way it interacts with other objects. The possible particle interactions are an necessary part of the particle's essence: we cannot leave them out.

Objection:
So far the discussion has just revolved around momentum states. However, we don't observe particles of fixed energy and momentum, but only at fixed points in space and time. Thus the construction is artificial, unintuitive and difficult to relate to common experience. •

Response:
I have stated things in terms of momentum creation state operators because it is easier to perform the calculations in that representation, since the momentum states are the meta-stable states of the system. It is possible to do a similar calculation in terms of operators that create or annihilate particles at particular locations – it is just a different choice of basis. However, when we want to compute things, we usually rotate the states to the energy/momentum basis, perform the calculations, and then rotate back to the space/time basis to express the result. Thus we need to consider the momentum basis in any case, so I have just been describing that, and will continue to focus on it.

---

physical world. The different renormalisation schemes are different but equivalent algebraic representations of the same physical reality, and in this sense we can say that there is a single solution.

However, it is still worth taking a brief look at the space-time formulation, for precisely the reason outlined in the objection: this has the more direct connection to what we observe and is more obviously related to our common perceptions. I will just look at the time evolution of Fermion fields in quantum electrodynamics, but the same process can be applied for the other fields and theories.

First of all, we need to define creation and annihilation operators for particles at a particular location, $\vec{\mathbf{x}}$. I will call then $\hat{a}^\dagger_{\vec{\mathbf{x}}}$ and $\hat{a}_{\vec{\mathbf{x}}}$ for the Fermion fields, and $\hat{\vec{\mathbf{A}}}_{\vec{\mathbf{x}}}$ for the gauge field. The equivalent of equation (338.1) is then

$$|\psi(t_0+\Delta t)\rangle = T[e^{i\Delta\iota\int d^3x \hat{a}^\dagger_{\vec{\mathbf{x}}}\beta(\vec{\alpha}\cdot(\vec{\nabla}-Q\hat{\vec{\mathbf{A}}}_{\vec{\mathbf{x}}})+\beta m+Q\hat{A}_{0,\vec{\mathbf{x}}})\hat{a}_{\vec{\mathbf{x}}}}]\,|\psi(t_0)\rangle\,. \tag{348.1}$$

So here we see three terms in the time evolution operator. The first is the mass term, $\hat{a}^\dagger_{\vec{\mathbf{x}}} m \hat{a}_{\vec{\mathbf{x}}}$. This connects the creation and annihilation operators at the same location, so it generates an amplitude that the particle will stay still. The greater the mass, the harder it is to get the particle to move, and this ultimately leads to the classical mechanics understanding of mass as something that measures how much force you need to get a given acceleration.

Then we have the interaction terms, $\hat{a}^\dagger_{\vec{\mathbf{x}}}\beta\vec{\alpha}\cdot\hat{\vec{\mathbf{A}}}_{\vec{\mathbf{x}}}\hat{a}_{\vec{\mathbf{x}}}$ and $\hat{a}^\dagger_{\vec{\mathbf{x}}}\beta\hat{A}_{0,\vec{\mathbf{x}}}\hat{a}_{\vec{\mathbf{x}}}$. These represent the amplitude that the Fermion will emit or absorb a photon. So this gives us the amplitude that the particle will interact with an external object, in this case a beam of light.

Finally, we have the kinetic term, $\hat{a}^\dagger_{\vec{\mathbf{x}}}\beta\vec{\alpha}\cdot\vec{\nabla}\hat{a}_{\vec{\mathbf{x}}}$. This is a little bit different, because the derivative operator connects the annihilation operator at one location with the creation operator at a neighbouring point in space-time. It provides us with the amplitude that the particle will move. This is a natural motion, which could be in any direction.

So we see three types of motion described by the Hamiltonian: staying still, interactions with external beings, and natural motion to the next place. These are analogous to the three types of motion allowed for in Aristotelian metaphysics (definition 4.5.5). What we see then is that the Hamiltonian operator represents a fight between these three types of motion. In practice, one of them will win at any given moment in time: the particle will either stay still, interact, or move to a neighbouring point in space. We cannot predict which one will win, but we can compute the likelihood of each event occurring. That's what the Hamiltonian operator does for us. We then sum up over every possible particle at every possible location, and apply the operator to a given input state to allow us to make predictions about what will happen next. ○

What I have described is (baring several minor simplifications) the theory of quantum electrodynamics. We have to incorporate creation and annihilation operators for each type of Fermion we observe (electrons, muons, protons quarks and do on), and give them the observed masses and couplings with the photon field. Combine this theory with statistical physics (using the quantum electrodynamics definition of the energy in terms of creation and annihilation operators rather than the classical definition, but otherwise closely based on Maxwell and Boltzmann's original work) and with gravity, and one obtains a theory describing almost everything we see around us. Optics, the physics of materials, electronics, atomic physics, chemistry: I have just written down a theory that fully describes them all. Add in gravity, and we have biology and the various atmospheric and earth sciences. Almost everything we see around us is described by the Hamiltonian (339.1) plus the principles of quantum field theory. It might seem a little cumbersome in places, but given that explanatory power, it can afford to be.

The theory has been checked to an incredible accuracy, with numerous different tests, for several different qualities based on wholly different experimental processes, agreeing within theoretical and experimental errors of between one to ten parts in a billion, depending on the quantity being measured. This is equivalent to computing the distance between the topmost point of St Paul's Cathedral in London and the topmost point of St Peter's in

Rome to the nearest millimetre (and, course, this is just the upper bound on the precision based on our most precise calculations and experiments – the theory will almost certainly continue to stand up as we improve the precision. It will almost certainly break down at some point, and need modification to include the effects of gravity and any physics we don't yet know about, but our experiments are probably still a long way from being precise enough to measure those deviations). I know of no experiment where the theory of quantum electrodynamics fails.

10.5. The Standard Model

That is electromagnetism. What of the other forces in nature? The basic idea is similar. For the weak nuclear interaction (the electro-weak theory), we assert that the spinors have an additional index with two components, so instead of a creation operator $a^\dagger$ we have $a^\dagger_w$, with the probability amplitude given by $\sum_{w=1}^{2} \langle 0| \gamma_0 a_{\sigma w} a^\dagger_{\sigma w} |0\rangle$. This means that the symmetry group is enlarged. Before, all we could do was rotate the complex phase of the field at each space time point, and this symmetry, given the name $U(1)$, led to electromagnetism. Now we have more freedom: we can rotate between the w indices as well as the complex phase for each index independently. The symmetry is enlarged; the new group is called $SU(2) \otimes U(1)$. $SU(2)$ represents three rotations around different axes; it is (equivalent to) the rotational symmetry group of the surface of a four dimensional sphere (just as the U(1) group is equivalent to the rotational symmetry of a circle). This symmetry can be expressed as a 3×3 real matrix, or a 2×2 complex matrix; the $SU(2)$ notation indicates that we are using the 2×2 complex representation.

The treatment of this local symmetry is almost exactly what we did in electromagnetism – we introduce a gauge field, select some transformation rules on this field to ensure that the action is gauge invariant, and so on. The only difference (for now, I will discuss another difference later) is that the gauge field is now a member of the adjoint representation of $SU(2) \otimes U(1)$. It is described by four independent parameters, so we now have four force carriers: as well as the photon, we have three more gauge Bosons, the W^+, W^- and Z^0, where the index indicates their electric charge. The most important difference from electromagnetism is that these new gauge Bosons can interact with each other, so there are additional vertices with three or four incoming gauge Bosons, which means that there are even more Feynman diagrams to consider. But the underlying principle of the expansion is exactly the same.

For the strong nuclear force (described by quantum chromodynamics, or QCD), we add a third index, c, to some of the Fermion fields (only the quark fields see this force), so we have $a^\dagger_{\sigma wc}$, with $c = 1, 2, 3$. c is unfortunately called colour. This name is, in my view, utterly stupid – it has nothing whatsoever to do with the colours we see around us - but unfortunately the earlier pioneers of the quark model chose that name, and since they were better physicists than I am, I have to fall into line and use their terminology. We introduce another set of gauge fields, which have the misfortune of being named gluons, which this time are members of the adjoint representation of SU(3) (a complex unitary 3×3 matrix with determinant one, which gives us eight gluon fields). This time there is a complication, in that the coupling (α_s, the equivalent of e in QED) is large at small re-normalisation scales, so the Taylor series expansion of the exponential of the Hamiltonian does not converge quickly enough that we can discount the larger terms in the expansion[20]; however at large re-normalisation scales the coupling is smaller and the Taylor series converges. That, then, would be the standard model of particle physics - if nature had been kind to us theoretical physicists.

Table 1 lists the fundamental particles of the standard model. There are four basic types of fundamental matter particle: the electron, the neutrino, the up quark and the down quark, together with their anti-particles. The electron and the neutrino don't have

[20]We are therefore forced to use other methods to calculate QCD observables at these scales.

Fermions:

Family	Lepton ($Y=-\frac{1}{2}$, $T^3=-\frac{1}{2}$, $g=0$)	Lepton ($Y=-\frac{1}{2}$, $T^3=\frac{1}{2}$, $g=0$)
1	electron (e)	electron neutrino (ν_e)
2	muon (μ)	μ-neutrino (ν_μ)
3	tau (τ)	τ-neutrino (ν_τ)
Family	quark ($Y=\frac{1}{6}$, $T^3=\frac{1}{2}$ $g=1$)	quark ($Y=\frac{1}{6}$, $T^3=-\frac{1}{2}$, $g=1$)
1	up (u)	down (d)
2	charm (c)	strange (s)
3	top (t)	bottom (b)

Bosons:

Electromagnetism	Weak Nuclear force	Strong Nuclear Force	Higgs
Photon (γ)	Z^0	gluon 1	Higgs Boson
	W^+	gluon 2	??
	W^-	gluon 3	
		gluon 4	
		gluon 5	
		gluon 6	
		gluon 7	
		gluon 8	

TABLE 1. The fundamental particles of physics; T^3 describes the coupling for left handed particles to one of the weak interaction fields and Y the coupling to the $U(1)$ field; in practice the photons and weak interaction particles are linear combinations of these fields; the electric charge is $Y+T^3$, g is the coupling to gluons. The quarks have three different colour states, so there are three different types of each quark – there are in practice 18 rather than 6 quarks. Each of the Fermions also comes with its anti-particle, which has the same mass but opposite electric charge.

colour, so they only interact with the electromagnetic and weak nuclear fields. The up and down quarks also have colour, so interact with gluons as well. This pattern of four different types of particle (and their specified electric charges) is actually not something arbitrary, but is required to make the electro-weak interaction self-consistent. We cannot, however, compute their masses from the theory: these have to be measured from experiment. The other surprise that there is not just one set of these four particles, but three. The whole pattern is repeated twice more at higher masses. Why three generations of particles, rather than just one or four? As far as I know, nobody has any idea of why this is the case. We can, however, be almost completely sure that there are three and only three generations, because the neutrinos are sufficiently light that if there was another one we would almost certainly have seen evidence for it.

Unfortunately, nature has a couple of nasty surprises up its sleeve. The first is that I have treated the masses of the Fermion fields as just a constant for each field. We can treat the total Hamiltonian of all of the particles as the sum of the Hamiltonians of the individual particles, so $\hat{H}=\sum_{i,j\in\text{Fermion field}}(a^i_{\sigma,w,c})^\dagger(\vec{\alpha}\cdot(i\vec{\partial}-e\vec{\mathbf{A}})+\beta M^{(ij)}+A_0)a^{(j)}_{\sigma,w,c}$, where M^{ij} is a matrix containing the Fermion masses. The Fermion fields we measure are the eigenvectors of this mass matrix, so in this natural basis we can treat the mass matrix as diagonal with the elements being the Fermion masses. The problem is that the eigenvectors of the mass matrix are slightly different in the electro-weak and QCD parts of the standard model, even though the eigenvalues of the mass matrix remain the same. This means that the creation and annihilation operators for the quark fields in QCD are in practice a linear superposition

of the quark creation and annihilation operators used in the electro-weak model (there is, of course, no fundamental reason why they should be the same); this is a bit of a pain, but can be corrected for relatively easily.

The second surprise is that the electro-weak theory does not work as I described it above. It is not simply an $SU(2)$ analogue of QED and QCD, although it is still an $SU(2)$ gauge theory. One point I have not yet discussed is that it is possible to project the massless theory into two states known (again, the name is stupid and misleading, but we are stuck with it) as left and right handed Fermions. The Lagrangian resembles

$$\mathcal{L} = \hat{\psi}^\dagger \gamma_0 \gamma_\mu (i\partial^\mu + eA^\mu)\hat{\psi}, \tag{351.1}$$

with $\hat{\psi}^\dagger$ and $\hat{\psi}$ the four dimensional analogue of the Fermion creation and annihilation operators introduced earlier[21], and we define

$$\begin{aligned} \hat{\bar{\psi}}_L &= \hat{\psi}^\dagger \gamma_0 \frac{1}{2}(1+\gamma_5) & \hat{\bar{\psi}}_R &= \hat{\psi}^\dagger \gamma_0 \frac{1}{2}(1-\gamma_5) \\ \hat{\psi}_L &= \frac{1}{2}(1-\gamma_5)\hat{\psi} & \hat{\psi}_R &= \frac{1}{2}(1+\gamma_5)\hat{\psi}, \end{aligned} \tag{351.2}$$

with $\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3$ anti-commuting with the other γ matrices, and $\gamma_5^2 = 1$. We can write the Lagrangian as

$$\mathcal{L} = \hat{\bar{\psi}}_L \gamma_\mu (i\partial^\mu + eA^\mu)\hat{\psi}_L + \hat{\bar{\psi}}_R \gamma_\mu (i\partial^\mu + eA^\mu)\hat{\psi}_R. \tag{351.3}$$

$\hat{\psi}_L$ is known as a left handed particle, and $\hat{\psi}_R$ as a right handed particle. In the massless theory, left and right handed particles can be treated entirely independently. The first problem is that the weak interaction only sees left handed particles – it does not interact with right handed particles. This is easily solved: we set the left and right handed particles as independent and give them different couplings (analogues to the electric charge) with respect to the weak fields. The second problem is that the Fermions have masses – and this whole decomposition forbids Fermion masses. The natural way of incorporating a Fermion mass, add a term $\hat{\bar{\psi}}_L m \hat{\psi}_R + \hat{\bar{\psi}}_R m \hat{\psi}_L$, would lead to a mixing between left and right handed particles - left handed particles would spontaneously become right handed, and vice versa, which would lead to observed interactions between the weak gauge Bosons and right handed fields, which are not seen. Thirdly, the theory I have been describing only makes sense if the gauge Bosons are massless, but the $W^\pm$ and Z^0 are observed to have masses.

So is the theory wrong? We should rather say that it is incomplete. We need one more actor on the stage, and then it agrees with reality and passes the test of experiment. The solution to this dilemma, introduced independently by several researchers, is to introduce another, spin zero complex field $\hat{\phi}$ with two components, with a Lagrangian

$$\begin{aligned} \mathcal{L}_{Higgs} = \hat{\phi}^\dagger(\partial_\mu - ie\hat{A}_\mu)^2\hat{\phi} + \mu^2\hat{\phi}^\dagger\hat{\phi} - \lambda(\hat{\phi}^\dagger\hat{\phi})^2 - \lambda_e \hat{\bar{\Psi}}_L \hat{\phi}\hat{\psi}_R - \lambda_e \hat{\bar{\psi}}_R \hat{\phi}^\dagger \hat{\Psi}_L - \\ \lambda_d \hat{\bar{Q}}_L \hat{\phi} \hat{d}_R - \lambda_u \hat{\phi}^\dagger \hat{\bar{Q}}'_L \hat{u}_R - \lambda_d \hat{\bar{d}}_R \hat{\phi}^\dagger Q_L \hat{\phi} - \lambda_u \hat{\bar{u}}_R \hat{Q}'_L \hat{\phi}, \end{aligned} \tag{351.4}$$

where $\hat{A}_\mu$ here represents the electro-weak gauge Bosons. In this way, the interaction between left and right handed Fermions is mediated by the particle $\hat{\phi}$. $\hat{\Psi}_L = (\hat{e}_L, \hat{\nu}_{e,L})$ gives the left handed leptons and $\hat{\bar{Q}}_L = (\hat{u}_L, \hat{d}_L)$ the left handed quarks with $\hat{Q}'_L = (\hat{d}_L, \hat{u}_L)$. This Lagrangian maintains all the symmetries we need for the theory, and is perfectly valid, if a little ugly.

The key point is that for suitably chosen couplings λ and μ the minimum energy state of $\hat{\phi}$ is not at zero, but at some non-zero value, say $\hat{\phi}_0 = (0, v)$ with $v = (\mu^2/\lambda)^{1/2}$. This means that, $|0\rangle$, the lowest energy state of the system, contains not an absence of these

[21] $\hat{a}^\dagger$ and $\hat{a}$ are the language used in the temporal, Hamiltonian, picture of quantum field theory; I have now switched to discussing the four dimensional, Lagrangian, picture, and a slightly different notation is used.

$\hat{\phi}$ fields, but specifically these fields in one particular state. This ground state could in principle be any complex vector with two components with a norm of v; but in practice one and only one of these must be chosen by nature. This is an example of what is known as spontaneous symmetry breaking, where a symmetry apparent in the Lagrangian is broken because it has multiple ground states mapped to each other by the symmetry but where each ground state itself violates the symmetry. Nature has to choose one of these ground states, and when it does the symmetry is broken. Since it doesn't affect any results which of these possible ground states of the field $\hat{\phi}$ is chosen, we may as well pick $\hat{\phi}_0$. Since we use the lowest energy states as the basis for our expansion, we should re-express the $\hat{\phi}$ field in terms of the deviation from $\hat{\phi}_0$.

We therefore rewrite $\hat{\phi} = \hat{\phi}_0 + (\hat{\phi}_1 + i\hat{\phi}_2)$, where the real fields $\hat{\phi}_1$ and $\hat{\phi}_2$ each represent two of the four dynamical particles contained in the field $\hat{\phi}$ ($\hat{\phi}_1$ and $\hat{\phi}_2$ each represent two component vectors, and each component represents an individual particle creation or annihilation operator). We assign creation and annihilation operators for these four particles, construct the Hamiltonian, and can include these particles in our Feynman rules and diagrammatic analysis. There are a few features to note.

Firstly, after we work through the algebra we find that three of our four particles are massless. This is a general feature of spontaneously broken symmetries – we find a massless particle for each symmetry we break. This is unfortunate. There is no massless spin zero particle in nature – we would have certainly seen it if it were there.

However, in this case we are saved by a mathematical curiosity. We can perform a gauge transformation that eliminates the fields representing the massless particles. These particles are absorbed into the gauge field – they are sitting inside the weak interaction gauge Bosons which we do observe; these particles are, like gauge symmetry, just an artefact of our choice of mathematical representation of reality, and are eliminated when we map back from the abstraction to reality. So that we don't see this particle in practice is in full agreement with the theory.

Secondly, a single representative of our original four scalar fields remains, and obtains a large mass. This particle is known as the Higgs Boson, and its existence (now experimentally confirmed) was a prediction of this theory.

Thirdly, the gauge Bosons and Fermion fields gain a mass. A mass term in the Lagrangian is a term proportional to the square of the field with no other fields or derivatives. When we set $\hat{\phi} = \hat{\phi}_0$, we create terms quadratic in both A and the Fermion fields proportional to the parameter v contained within the definition of $\hat{\phi}_0$. In quantum field theory, a quadratic term in the Hamiltonian (or Lagrangian) for a Boson field without any derivatives represents a particle mass (you can see this in the case of the Fermion Hamiltonian in equation (338.1)). In other words, the observed Fermion and gauge Bosons behave as though they are massive, there are independent left and right handed fields, and yet the full set of symmetries is still maintained, at least at the level of the Lagrangian before symmetry breaking.

Now, the electro-weak symmetry breaking (as it is known) need not take precisely this form (the potential, for example, need not be as I have written it), so we do not have a precise theoretical prediction for the Higgs Boson mass, or even know for certain that there is only one Higgs Boson. There are various different models each consistent with the experimental requirements; what I described above was only the simplest. But we do know that the consistency of the theory of the weak interaction requires that there is something out there. There must be a massless particle absorbed into the gauge field, and at least one more particle beyond the Fermions and gauge Bosons. But combining this $\hat{\phi}$ with the previous discussion of the weak interaction – to obtain the Glashow, Weinberg and Salam electroweak theory, gives us a theory which at least explains the dominant features of the weak interaction – it is an $SU(2)$ gauge theory which only couples to left handed Fermions, where the Fermions and gauge Bosons are massive. But how does it fare against experiment

when it comes down to the details? Wonderfully well. Like Quantum electrodynamics, it has passed the experimental test to an incredible precision.

So that is the standard model, the most successful physical theory yet devised by man. It agrees with experiment to amazing precision, and we only know of one place where it breaks down - it predicts that the neutrinos are massless, when in practice they have small masses -- that, however, only requires a minor modification to fix. The question is not whether it is possible to adjust the standard model to account for neutrino masses, but which particular hypothesised addition is correct. That is something which will have to be determined by experiment. Not only that, but QCD and QED are extremely beautiful theories, with an immense range of physical results and breathtaking mathematical structure proceeding from a small set of axioms. The weak interaction/Higgs sector is (in my opinion) a complete mess, but the ultimate judge is experiment, and nature seems to like it even if I don't.

Of course, there are outstanding problems, both in calculating the predictions of the theory to compare against experiment - most particularly low energy QCD, where the procedure of expanding the bare theory in terms of diagrams of ever decreasing importance and then renormalizing breaks down, although even here good progress is being made using numerical calculations -- but also at a more abstract level. Why are there three families of Fermions? Are there any other particles at higher masses (perhaps Boson fields)? Why an $SU(3)\otimes SU(2)\otimes U(1)$ gauge symmetry? (Is it only an $SU(3)\otimes SU(2)\otimes U(1)$ gauge symmetry?) Why is the weak sector so different in structure from QED and QCD? How do we incorporate gravity into this theory? Why are the masses and couplings the values that they are (in the theory, the Fermion masses and couplings with the gauge Bosons are, in effect, free parameters -- they could take any value, but they do take certain values; so why those particular values)? Why do we calculate amplitudes using the naive vacuum state, i.e. the state with no particles[22]?

The standard model is not the full theory of nature. Aside from neutrino masses, the standard model has stood for thirty-five years, passing experimental tests to a very high accuracy, and we are no nearer finding what lies beyond it. Something does -- gravity needs to be merged with the quantum field theory picture in a mathematically consistent way, and this has not yet been done (or rather, the various proposals are still far from being confirmed or ruled out by experiment) -- but what that is is as yet unknown. Nonetheless, the success of the standard model means that we can be as certain as we can be about anything that the fundamental physical theory of nature is a quantum field theory, built along similar lines and similar principles to quantum electrodynamics. Thus we are now ready to start from what is known about quantum field theories and build a philosophy of physics fit for the twenty first century.

[22]I rarely see this question asked, but in my view it is hugely significant. It is well known, for example, that the vacuum gauge field in non-Abelian theories (QCD and the weak interaction) is not naturally flat, but contains distinct structures. Similarly, the vacuum state for the Fermion fields is likely to contain high momentum/short distance fluctuations of Fermion and anti-Fermion fields. This will forbid the creation of high momentum 'virtual' Fermions, and provide a natural cut-off for the theory.

CHAPTER 11

The Philosophical Implications of Quantum Field Theory

11.1. Form and The Physics of Composite Beings

DEFINITION 11.1.1. **Particle**:
A particle is a meta-stable state which carries a definite energy and momentum, and which corresponds to a particular set of creation and annihilation operators. That it is meta-stable means that the creation and annihilation operators must commute with a non-interacting Hamiltonian (the time evolution operator), with the interaction part of the Hamiltonian sufficiently small in comparison that the transition rate of the particle (governed by the interaction constant via Fermi's golden rule) is greater than the inverse Planck time (Energy divided by $\hbar$) of the particle. A fundamental particle is one which cannot be treated as a bound state of other particles. A particle is thus an energy/momentum excitation of the underlying matter field.

Note that this definition is contrary to the classical definition of a particle, which implies an element of spatial localisation. In the quantum theory, if the momentum state is definitely defined, then the location of the particle is wholly undefined, so discussing localisation makes no sense. Nor can quantum particles be thought of in terms of classical fields. A classical particle is localised and discrete; a classical field de-localised and continuous; but a (non-interacting) quantum particle is both de-localised and discrete[1]. In our classical way of thinking about the world this combination is difficult to imagine; however, that's just another limitation of our imagination: we naturally imagine that space and time are more fundamental, while quantum particles are discrete and 'localised' in the space of momentum/energy. Quantum particles do not resemble any classical objects, and the categories developed from classical mechanics are simply not relevant. We need something new. This is related to the very different metaphysics that lies behind classical mechanics and quantum physics, and in particular the point that in classical mechanics both location and momentum are needed to specify a physical state, while in quantum physics momentum alone is sufficient, and both momentum and location together are impossible. A particle is defined in terms of the momentum because only momenta states commute with the free Hamiltonian and thus are meta-stable in time. Technically, therefore, we should discuss 'classical particles' or 'quantum particles' to make this distinction clear, although I will usually drop the modifier because the context makes it obvious which one I am discussing; for almost all of this work (except when specifically discussing classical mechanics) when I discuss a particle, I mean a quantum particle. Δ

Something similar to the standard model is almost certainly the fundamental theory of nature. But this deals with particles at their most basic level -- the ultimate building blocks of matter. How do we get from this to the complex composite structures we see around us?

[1]That is a free (non-interacting) quantum particle: interactions force the space time particle probability amplitudes to peak around certain locations, which is why in classical physics particles have the appearance of being localised, as long as you don't look too closely.

Even the atom is difficult to get to from the fundamental physics, once we go past Hydrogen and Helium. We know how it is done in principle the up and down quarks are bound together by gluon exchange to make protons and neutrons; protons and neutrons are bound together also by gluon exchange to create the atomic nucleus; this creates an electromagnetic field with its associated Energy states, and electrons sit in the eigenstates of this Hamiltonian. The quarks themselves can interact with weak force, which converts up quarks into down quarks plus a W Boson, which then decays into an electron and neutrino (one form of radioactive decay). Then the atoms bind together into molecules or crystals. In the case of molecules, the two atoms come together, and their outermost electron energy levels are merged together, allowing the final group of electrons to flow between both atoms and hold them together. In the case of metallic substances, we once again see the highest energy levels of each atom interacting and influencing each other, only this time instead of being localised around only a pair of atoms, they come to extend across the whole metal, allowing electrons to easily pass from one atom to another.

Each step of this process is understood, at least in many cases, and we have no reason to suppose that the cases we don't understand are any different in principle from those we do. The problem is not the underlying principle, just the difficulty of the calculations. In some ways condensed matter physics, the branch of physics that studies solids, is very different from particle physics – in this case the fundamental theory, quantum electrodynamics is known, but the problem is just the huge variations of circumstances in which it has to be applied. However, in other ways it is very similar to particle physics, because the underlying mathematical structure is very similar. The Maxwell-Boltzmann distribution for a system of many particles is an exponential of the energy. When we move to the quantum analogue of classical physics, this just becomes the exponential of the energy sandwiched between creation and annihilation operators. The only difference between this and the mathematical formalism of the quantum field theory (equation (339.2) is a factor of the square root of minus one. Condensed matter physicists usually use the non-relativistic limit of field theory, because relativistic effects are not so important for their studies and its slightly easier, but they still need to deal with the creation and annihilation of particles. Modern condensed matter physics uses Feynman diagrams as much as particle physics, just with different rules.

One important concept in modern physics is what is known as effective field theory (Burgess, 2007; Leutwyler, 1994; Atland and Simons, 2010; Shankar, 1997; Schakel, 2008). This is related to our freedom to rotate creation and annihilation operators into another basis, but takes it in a different direction. Instead it combines creation and annihilation operators for different particles. In my own field of particle physics, for example, we know that two quarks can be bound together into a particle known as a Meson. What we have in the fundamental theory is two quark creation operators. By a mathematical trick, we can re-write the theory in terms of a Meson creation operator and various other terms. We can then use the Meson creation and annihilation operator to develop a quantum field theory describing Mesons and their interactions. We know roughly what form the Hamiltonian operator for Mesons must take because of symmetry considerations (and it is considerably more complicated than the standard model Hamiltonian), so we can write down that Hamiltonian, use it to construct the Feynman rules for a system of interacting Mesons, and get calculating. Three quarks make up a Baryon, a proton, neutron, or various other particles. We can do the same thing. We use a rotation of the coordinate basis to combine the quark and gluon creation and annihilation operators into the Baryon creation and annihilation operators (plus other states), use symmetries to constrain the Hamiltonian, and that gives us a theory which describes the interaction of protons and neutrons – i.e. nuclear physics. Though in principle possible, the actual steps to construct these theories from first principles are very difficult (in the sense that something which is to all practical purposes impossible is difficult), but fortunately we know roughly what the final result must resemble to be consistent with the underlying symmetries, and with the help of experiment

to compute the various missing parameters, we can construct the underlying theory. All of these theories, of course, are quantum field theories.

This can be extended to systems of solids, and once again the use of effective field theories allows us to map out the energy levels of the various materials. The most important players are the electrons, photons, and also new 'virtual' particles called phonons[2], which represent the central core of the atoms and in particular their wave-like vibrations. These energy levels allow us to compute the various properties of the materials, such as their colour and conductivity. For example, in insulators, the energy levels (which are each a function of the electron momentum) are formed into discrete bands. Within each band, the energy levels are tightly packed, so an electron can easily hop from one energy level to another. However, there is a gap between the bands, making it hard for an electron to pass from one band to another. Generally speaking, the lowest energy levels tend to be localised around the atom or a cluster of atoms, while the higher energy levels might be spread out across the whole substance, meaning that an electron in one of these levels has a non-zero probability of being found anywhere in the material. These energy levels can be found from the eigenstates of the Hamiltonian, which is now built from the creation and annihilation operators of all the particles within the atoms of the substance.

For an insulator in its lowest energy state, the low lying energy levels are fully occupied, and then there is a gap until the next energy level. It is not possible for an electron to be excited to an unoccupied energy level except by the emission or absorption of a photon with a reasonably large energy. The allowed energies of these photons are given by the separation between these energy bands. Thus only certain photon energies, which correspond to the colours of the material, can be absorbed and then re-emitted (i.e. reflection or scattering). Other colours correspond to photon wavelengths which would seek to excite an electron to an energy where there gap in the eigenstates of the Hamiltonian. The gap means that no electron in the substance can have that energy, and therefore that particular colour will not be reflected by the material. This gives rise to the many different colours of electrical insulators. Metals (in general) do not have a gap between the energy bands. This allows electrons to flow freely around the material, allowing electrical conductivity, and also means that every photon energy can in principle be absorbed and then re-emitted by the material, meaning that metals can reflect every colour of light (though not necessarily evenly, as the various coloured metals such as gold show).

Colour is an aspect that many philosophers ponder over. There are actually two aspects of colour; the first is related to the materials themselves, and the second to how the sensations are processed in the human mind to give our perception of colour. Confusion between these two aspects is unfortunately far too common. The question of most interest to me as a physicist is how colour is present in the material. What does a red bus and a red postbox have in common that makes us both perceive them as red? The answer is that their surfaces both have a similar structure of energy bands. But the energy bands can be computed from the underlying Hamiltonian. The same is true for any other observable property of the material: ultimately it can be computed from knowledge of the Hamiltonian of the effective field theory describing that material, expressed in terms of creation and annihilation operators which describe the possible interactions of the substance.

In summary, rearranging fundamental quantum field theory leads to an effective field theory, which gives rise to a Hamiltonian for the more complex substance, and the discrete energy levels of this Hamiltonian, together with its interactions with external gauge Bosons (particularly photons) give rise to the observable and macroscopic properties of a substance.

[2]Phonons are an artefact of the remapping to the appropriate representation of the creation operators which is useful for studying solids. When we rotate the basis, we end up with new creation operators representing the electrons and photons, but have some left over which describe additional degrees of freedom of the system. They are not real particles as electrons and photons are, but creation and annihilation operators that describe transitions between various vibrational states of atoms. They enter the equations in the same way as an actual particle. So they are treated as though they were particles.

The complete answer of how properties relate to being might be more than the physical account, but it certainly cannot be less than it, nor can it be inconsistent with it. I discussed colour and electrical conductivity, but I could have equally discussed rigidity, or brittleness, or transparency, or how sound waves are propagated, and a host of other properties. All of these depend on the effective Hamiltonian of the substance, and how it interacts with electrons, photons, and the virtual phonons.

The Hamiltonian of an effective theory describing a more complex substance is not simply the sum of the Hamiltonians of the fundamental particles that constitute that substance. The process of joining together various creation and annihilation operators is not a simple linear relation. What this means is that the energy eigenstates of (say) a water molecule are not just the eigenstates of the two hydrogen atoms and the oxygen atom added together. It is something very new and different, with entirely different properties. Water is not just hydrogen plus oxygen. Water is something different, with the energy levels of the atoms modified unrecognisably by the interactions between the atoms. Equally, a speck of gold is not just a thousand billion billion gold atoms; its energy levels are quite different from the atoms in isolation. The energy levels combine in such a way that new structure and new properties arise.

For example, phonon interactions can be strong enough to bind two electrons together into a molecule, overcoming the electromagnetic repulsion between them. But two spin half particles joined together is no longer a Fermion, but a spin 0 or spin 1 Boson. With Fermions, we can only have one electron in each energy eigenstate, so a moving electron has to overcome an obstacle course, navigating from one energy level to another to avoid being in the same eigenstate as another electron. This process takes energy, which causes electrical resistance. But pairs of electrons are Bosons: as many of them as you like can occupy a single energy level. All the Cooper pairs (as they are called) can exist simultaneously in the same energy state: they can pass through each other, and that means electrical conductivity without resistance. There are also new geometrical and topological effects in some substances which can force the conductance to take discrete values. By taking advantage of these various features, we have modern electronics.

The energy eigenvalues and eigenstates thus contain all the information about the possible states a substance can be in. They are not a single state, but a set of numerous possible states. All the observable properties of the substance can be computed from them (in principle all can be, and in practice many properties of many substances are fully understood in terms of the energy eigenstates and band structure). They do not tell us which state a substance is in at any given moment of time, but they do tell us the possible states and the possible transitions between states.

The Hamiltonian can be distorted still leaving us with a stable or meta-stable substance, and in practice often is, either because of external fields or temperatures, or because the particles might move around. However, we can distinguish between continuous distortions and discontinuous distortions of the Hamiltonian. It is not possible to get from every possible Hamiltonian to every other possible Hamiltonian via a continuous change of the energy levels. Instead, the Hamiltonians can be classified into various classes; each region internally connected to each other by continuous changes of the energy levels and jumps from one energy eigenstate to another, but disconnected from the other regions. In this way, the Hamiltonian also provides us with a convenient way to distinguish one substance from another.

So the effective Hamiltonian, or equivalently the set of continuously connected energy eigenstates and eigenvalues, is an abstract principle which is closely related to material objects, which contains a set of states describing the possible internal changes of the substance, and from which the properties of the substance can be computed and which distinguishes one type of substance from another. But these are precisely the properties which the scholastics

attributed to form; the only difference is that the Hamiltonian is a mathematically well-defined abstraction while the scholastic form was often left vague and undefined, at least in terms of the microscopic details. The scholastic philosophers knew that there must be something with these features; modern physicists have found it. Of course, the Hamiltonian alone is not enough to describe the material; one of the eigenstates or potentia (the two concepts can now be identified with each other) must be actualised, and there must be some principle telling us which state this is. In scholastic metaphysics, this principle is matter. The only difference is that quantum field theory allows for matter to exist in a superposition of the energy eigenstates rather than being forced into one alone (except when we take a measurement). But this is only a minor detail; the philosophy can be modified[3] without causing any real damage to the system as a whole by incorporating this idea.

I will therefore define the form as follows:

DEFINITION 11.1.2. **Form:**
The form of a material being can at least in part be represented by the various eigenstates of Hamiltonian operators (effective field theory Hamiltonians for complex systems) which are connected by continuous deformations; with these connections only passing through meta-stable states. Each substance is described by a different region of continuously connected Hamiltonians. Δ

Discussion:
Note that the continuity refers to changes in the Hamiltonian operator. The jump from one eigenstate to another usually implies a discontinuous deformation of the fields; the Hamiltonian remains unchanged. A continuous deformation of the Hamiltonian would also usually (not always) lead to a continuous deformation of its eigenstates, and thus the field configurations that make up the potentia.

Unlike the scholastic description of form earlier, this definition is certainly not circular, although it depends on the definition from physics and mathematics of the Hamiltonian operator and eigenstates. Different substances are clearly distinguished from each other by regions of continuity.

One caveat that needs to be considered is how to treat degenerate energy eigenstates, i.e. different physical states which have the same energy eigenvalue but correspond to operators which each commute with the Hamiltonian but not with each other. As an example, consider the various spin operators. One can measure spin in any of the three spatial directions, and the spin operators in each direction do not commute with each other, so the choice of energy eigenstate depends on an arbitrary choice of spin direction. The best approach here is probably to note the degeneracy of the states, and list the operators that can distinguish them, and include that in the set of information represented by the form; for an actual state one would have to specify the spin operator used and its eigenvalue. For example, one could say that a particle has spin $\frac{1}{2}$ along the x-axis; with its possible spin states along (for example) the y-axis as other potentia, with a certain transition probability (or final cause) into each of those states. This is thus, I think, more of a complication than a difficulty, and I will subsequently just mention 'energy eigenstate' leaving these extra details implicit.

[3]For example, we can always rotate the basis of states so that the actual state is no longer a superposition, but in a single state in the new basis. For example, we can draw an x and y axis, and then a diagonal line. This line is neither wholly along the x axis, nor on the y axis, but a combination or superposition of the two. But the choice to draw the x axis horizontally and the y axis vertically is arbitrary. We could draw a new x axis so that it lies along the line, and the new y axis perpendicular to it, and in this new coordinate system the line is no longer a superposition of the two directions but directly along the x axis. This redrawing of the axes is a rotation of the coordinate basis. The potentia are not simply the eigenstates of one given representation of the Hamiltonian, but every state in every possible representation would be a different potentia.

Scholastic metaphysicists may complain that there is more to the form than just this mathematical abstraction. I am willing to concede the point for the sake of argument (though I cannot think of what else could contribute), and for this reason I added *at least in part* to my definition. However, everything that is in metaphysics derived from the form is, from the perspective of modern physics derived from the Hamiltonian operator. Likewise both the form and Hamiltonian describe various sets of potentia in the case of form and eigenstates in the case of the Hamiltonian. We can describe both how things can change and yet remain the same substance. In metaphysics, this is movement from one potentia to another; in physics movement from one energy eigenstate to another.

Note also that the Hamiltonian is constructed from creation and annihilation operators, and each creation and annihilation operator describes the allowed transitions between energy eigenstates (or potentia), and also how the substance might react with something external to it. The list of possible interactions, transitions and effects is contained within the effective Hamiltonian of the substance. This list of possible effects is, of course, no more than what the scholastic philosophers called the final cause of the substance. It can be computed from the structure of the various effective Hamiltonians. Equally, we can take a list of tendencies and properties, and construct the possible effective Hamiltonians consistent with that list. The scholastic philosophers postulated a connection between the formal and final causes of a substance. They knew that there had to be such a link, but had to leave it vague and undefined because their physics was inadequate. Once again, where the medieval philosophers stumbled, contemporary physicists have provided the answer.

Some care has to be taken to what is meant by 'continuously connected'. Mathematically this is well defined. However, from the perspective of metaphysics this mathematical definition doesn't quite agree with what we are accustomed to. It can however be extended to allow small discrete steps of fixed size. This gives rise to the concept of various levels of form. For example, we might have the strict idea of form of the mathematical continuously connected. We then might allow a slightly less rigorous definition by allowing certain types of change (the addition of another atom into the crystal, for example - a discontinuous change, but not one that changes the properties of the substance other than its mass) to allow a broader grouping together of similar rather than identical substances. Then we can allow for broader kinds still by allowing even more drastic changes.

Suppose, for example, we have some pure gold: nothing but gold atoms in there. Now to turn this into an alloy, we would need to turn some of the gold atoms into another type of atom. This means adding protons and neutrons from somewhere, and it certainly counts as a discontinuous change. Thus the alloy is not the same substance as the gold, which is all well and good because that is what we want. However, suppose that we add another gold atom to the edge of the crystal. This also counts as a discontinuous change, though not one that would usually affect how we identify the substance. We can overcome this by either saying that the form also describes precisely how many atoms there are in the substance, or by adopting several layers of form distinguished by various different types of change, so perhaps we don't count the extension of the energy levels (which are periodic) over a larger range of space as a discontinuous change for the sake of defining form. So by increasing the permitted size of the steps allowed to stay within the same kind, we can first of all have one particular speck of gold (mathematical continuity); to (being a little less strict by what we mean by continuous change) to gold in general; to alloys of gold; to metals; to all solids and finally all material beings.

Another caveat is in our treatment of living beings. I am in favour (unlike many more professional Aristotelian philosophers) of saying that each individual animal has its own form. Of course, there are issues here: is it the same man if you chop off one of his limbs? The natural answer is 'Yes', and we could of course duplicate this effect by removing one cell at a time. So we have to define 'continuous change of the Hamiltonian' in such a way that allows the addition and subtraction of cells to be counted as part of the same form

(which we have to, to describe the development of an individual from Zygote to adult, let alone all the minor adjustments to our body made during the course of our lives). It might be objected that the body is made of many discrete parts, and should not be treated as one substance, but these parts are interlocking and dependent on each other for their healthy functioning, so they should not be considered separately. It might also be objected that writing down a Hamiltonian operator for something as complex as a living organism is in practice impossible; however that we cannot do it does not mean that it cannot be done. We know that each individual part and interaction between parts can be treated quantum mechanically, and therefore the whole being ultimately is also made of quantum fields. We do not know what the solution is (and we would be foolish to try, since there are methods better suited to the biological sciences), but we do know that there is a solution, and that is enough for the metaphysics.

Allowing slightly larger steps between living organisms can lead us to identify species as the next level up. In traditional scholastic metaphysics, form applies at the level of the species rather than the individual; I am not enough of a traditionalist to see this as necessary, but it is nonetheless important to understand to what extent we can discuss 'the form of a tiger' or 'the form of an oak tree'. Each individual living organism is defined by a unique DNA pattern, coupled with the various experiences that determine precisely how they grow and develop from that initial genetic key, plus, of course, the underlying matter that distinguishes clones or twins. We also know that members of a species have closely connected DNA sequences. We also know that almost all DNA sequences lead to non-viable organisms, either those which die the moment they are conceived because they lack a liver or heart or some key organ needed for them to survive, or simply because, although the creature can in principle survive, in practice it is not strong enough to withstand the challenges of natural selection. Thus we can distinguish between viable and non-viable DNA sequences. We can also consider a set of sequences of DNA which differ only by the switching of one 'letter' or the addition or subtraction of single letters. We can in this way describe regions of DNA sequences which are internally connected by changes such as this which give viable creatures; and these regions will certainly be disconnected from other types of creature.

But doesn't evolution by natural selection say that all life is ultimately connected to each other by single simple changes to the DNA? Yes, but not all those intermediate steps are viable enough to survive natural selection in a competitive environment. We know that, because we do see genetic disconnects between different types of creatures, and if those intermediate steps were as competitive as the modern creatures, then we ought to see them alive today. If there is a global catastrophe, a sudden change in the environment, then that will mean that there will be a big shake up in the range of viable creatures; some which were well adapted to the old environment will not prosper in the new one, leading to a large scale extinction, and lack of selection pressure. In this case, those intermediate animals might be able to survive for long enough for them to tunnel through the barrier between species, allowing new species to emerge. But such genetic sequences should not be considered as viable because they are not meta-stable. This does not offer an obstacle to the idea that something half way between a man and a chimpanzee would not be able to compete against either of the fully evolved forms.

Of course, this definition of species will not be in full agreement with our common conceptions. There are, for example, ring species: in Western Europe, the herring gull and black backed gull are regarded as separate species. However, the genetic code gradually changes as we travel East. We start with black backed gulls, then as we move into Eastern Europe the genetics becomes slightly closer to the Herring gull, and across Russia slightly closer still, and into Alaska we cannot tell whether it is a black back gull or a Herring gull, then the North American Herring gull, then across the Atlantic to the European Herring gull. Each subspecies is genetically close to its neighbours, but as we transverse the world,

these small changes all add up to something substantial. Many other examples of 'speciation' are between creatures genetically close enough that the intermediate forms could be viable, but for one reason or another the genetic drift led the communities apart. By this classification, these gulls and sea snails and other organisms would for each type of animal have to be regarded as one species. This does not, however, show that there is a viable bird midway between a gull on one hand and a blue tit or an eagle on the other (if there were a continuous path, then it would destroy my suggestion), and thus I do not think that these counter examples are enough to overthrow my classification. But the details of this I will have to leave to those more expert in biology than I. My tentative feeling is that something along these lines will work, and then we can treat forms of species with a similar definition to forms of inanimate substances. □

11.2. Metaphysical Implications of Quantum Field Theory

One thing has always puzzled me while reading of the main approaches taken in contemporary metaphysics: that they make little reference to what physicists have to say about the same things they are discussing (Loux, 2006). So, for example, when discussing the relation between the properties of a being (such as colour, shape, texture and so on) and the being itself, the metaphysicists discuss whether the being is just a bundle of the properties, or the properties belong to a substratum beneath the being (since they cannot belong to the being itself, otherwise we could not explain how a being might lose a property and still be the same thing), or whether properties exist at all in a meaningful sense. What surprises me is that these discussions seem to operate in a vacuum entirely separate from physics, even though these are the very questions which physics attempts to answer. To repeat a point I made earlier, the complete answer of how properties relate to being might be more than the physical account, but it certainly cannot be less than it, nor can it be inconsistent with it. This is not to say that metaphysical systems developed before quantum mechanics became mature are useless: they may by chance be consistent with its principles; nor that those developed after it will be useful – their creators may have been ignorant of or misunderstood quantum theory – but each system must be judged against quantum theory, and if it is found to be inconsistent with it, discarded.

There are, of course, philosophers who do think about quantum physics, and they are to be commended. But even here most of them that I have read only consider quantum mechanics rather than quantum field theory (Jammer, 1974; Rae, 1986; Maudlin, 1994; Polkinghorne, 2007; Saunders et al., 2012; Baggott, 2003). While the two are related, they are not the same, and there are important differences. For example, quantum mechanics has no mechanism for the creation and annihilation of particles, while in field theory particles interact only through particle creation and annihilation. In other words, in quantum mechanics, particles cannot come in and out of existence, while in field theory every interaction is described by particles being created and destroyed. This has implications, most notably in my representation of final causality. In quantum mechanics, the evolution of the wavefunction is deterministic, while the measurement process is indeterminate; in field theory there is no determinism at all for single particles: the Fock state describes all particles in the physical system. Quantum mechanics is wrong. It is usefully wrong, in the same sense that the work of Copernicus was wrong (since his answer for solar system dynamics was, when we go down to the final details, incorrect: not even a good approximation in a certain limit of the correct theory); but it was nonetheless useful because it started people thinking along lines which eventually led to the 'right' answer.

In the same way, quantum mechanics, though it shares many features with field theory, and its answers for some problems are pretty close to getting the right answer, is an incorrect theory of nature. It is not even incorrect in the same sense as Newton's theory (mathematically equivalent to the $c \to \infty, \hbar \to 0$ and flat space-time limit of the correct

theory) or the standard model of particle physics (which is the low energy part of the flat space-time limit of the correct theory) – it is just irredeemably wrong if we are searching for the most fundamental physics. So I tear out my hair whenever I see a philosopher of physics discuss the Schrödinger equation in its non-relativistic or quantum mechanical form – for philosophers, who ought to be concerned only with the most fundamental physics, that equation is a distraction and an irrelevance (and ought to be mentioned by philosophers only for its historical importance and as a means of introducing the correct physical theory). The number of philosophers which use field theory as the starting point of their metaphysics is, unfortunately, minuscule; one reason why contemporary philosophy is generally heading in the wrong direction.

Of course, this doesn't mean that every proposed philosophy of quantum mechanics is invalid in field theory, and vice versa. There are some, such as the Everett multiple-worlds interpretation which carry forwards directly. But there are other interpretations of quantum mechanics which seem unlikely to be reconciled with field theory. The classic example of this is the de-Broglie/Bohm pilot wave interpretation of quantum mechanics, which is generally considered to be incompatible with relativity without taking some desperate measures such as backwards causation (Gisin, 2011; Drezet, 2019). At the very least, it will require considerable modification perhaps to the point of making it unreasonable[4]. It might also, of course, be possible that there are valid interpretations of quantum field theory which cannot be carried back to quantum mechanics. Quantum field theory is not so much more difficult than quantum mechanics to understand, at least until we get into the detailed calculation of observables. So my advice to philosophers of science would be to forget quantum mechanics, learn field theory, and use that as your starting point.

So what are the physical features of quantum field theory that might have implications for metaphysics? Firstly, the idea that the allowed states of the system are eigenvectors of the Hamiltonian operator, and that these can be discrete. This is different from classical physics, and means that the way we identify being should not focus on the geometrical features of location, extension and motion but instead these discontinuous states each represented by a different eigenfunction. In particular, whatever metaphysical system we use has got to focus on well defined and discrete states, that each of these states represents a different way a being can be of the same substance, and in particular be able to handle changes from one eigenstate to another. However, we also need to account for the idea of linear superposition; that something need not be in one definite state, but can (or so it seems) be in several at the same instant, or have its state unspecified between several states. In fact, this idea of discrete energy eigenstates is essential for the stability of matter, since otherwise without it matter would change with every small perturbation. Combined with the Pauli exclusion principle, chemistry would be impossible without this. Classical physics is not only wrong, it cannot lead to chemistry and thus all forms of life that we know of.

Another feature of quantum field theory which is alien to both classical and Greek physics is the idea that certain operators are non-commuting, i.e. you can't simultaneously measure both observables related to that operator. It is not only not necessary to invoke both momentum and location to specify the state of a particle, it is impossible to do so. Furthermore, Quantum field theory is probabilistic rather than deterministic, even more so than quantum mechanics because what is left of the wavefunction no longer evolves deterministically (at least in a particle interpretation of field theory). This means that any certain prediction of the future is impossible; we can, however, express things in terms of probabilities or probability amplitudes.

In particular, several contemporary metaphysical systems view the properties of something as key aspects of defining what that thing is; so an object is identified by a list of

[4]Of course, this hasn't stopped people from trying; see (Struyve, 2011) for a recent review. All of the approaches I am aware of have serious drawbacks.

predicates or properties. In quantum mechanics, however, the key element is the wavefunction, and in field theory the Fock state after being acted upon by some creation operator. The expected values of observables (which correspond to the properties used by the philosophers) are computed by applying the corresponding operator to the wavefunction or Fock state. The wavefunction or Fock state are fundamental; the properties secondary. Other properties, such as colour, depend on the energy differences between different eigenvalues of the Hamiltonian. Indeed, it is not possible to list every possible property of an object and tick them off as being either present or absent because the operators corresponding to the observables do not commute with each other: if the particle has a well-defined location, then its momentum is completely undetermined; if it has a well defined spin in one direction, its spins in the other directions are undetermined, yet in each case the particle's wavefunction, or the part of the Fock state corresponding to that particle, is perfectly well-defined. While it is tempting to use properties as the key parts of understanding the nature of objects, since observables are what we directly encounter, and in classical mechanics they are sufficient, can be known simultaneously, and there is no reason for particles to be indistinguishable. However, when we consider quantum physics, metaphysical approaches such as this should be avoided. I am not, of course, claiming that the wavefunction or Fock state is a complete description of the particle; it is just an abstraction representing that part of the particle which can be represented algebraically. However, it is nonetheless one step closer to an understanding of the particle's essence than the particle's observable properties, which are perforce even further from an expression of the nature of being as being.

Another key feature of the quantum field theories that describe nature is locality: there is no action between substances at a distance, either in space or in time.. All the interactions are point-like, composed of two fields or their differentials at the same location. This is true for Einstein's theory of gravity as well, which is just given as a straightforward set of differential equations, with no mention otherwise of the distance between objects; Newton's action at a distance emerges when we try to remove the intermediate step of the curvature of the metric tensor from the theory. In quantum field theory, locality plays an important role in that it guarantees the conservation of energy and momentum in each interaction, which is key to principles of causality. This condition is not strictly necessary for a quantum field theory; quantum field theories in discrete space time (so the only space time positions allowed are an integer multiple of a lattice spacing) use similar Lagrangians and are well defined quantum field theories. But such theories are only used as approximations to nature, and not seen as nature itself. And even here, one cannot have interactions of arbitrarily large separations. The interaction term between matter must decrease at least exponentially with the distance r.

There is, however, room for quantum events to be correlated in a non-local way. The locality applies to the Hamiltonian: the process of creation and annihilation; how one substance emerges and decays into another. This is substance causality. The Hamiltonian, however, only describes which processes are possible, not which ones happen in practice. To understand which of these things actually occurs, we need to add something extra to the Hamiltonian evolution of wavefunctions. This something extra "determines" quantum events. Since it is not described by the Hamiltonian, there is no need for it to be local. Substance causality is local; that is built into the physics. Nothing I have written so far implies that event causality (or, more precisely, correlation between events) should be local, and in practice there are occasions when it isn't (see section 15.5).

There are two reasons why we require locality in the Hamiltonian and thus substance causality. The first is that non-local operators are associated with operators which allow an infinite momentum. There is a theorem that states for any momentum operator, $\Gamma(p)$, that disappears rapidly enough at infinite momentum[5] that is not infinite along the real axis

[5]We extend the momentum p to the complex plane; strictly we require that it is zero at $|p| = \infty$ and not just $p = \pm\infty$

the corresponding space-time operator must decay at least exponentially, that is at least as $\Gamma(x,y) \sim e^{-|x-y|/a}$. a here represents some length scale. But there is no preferred length scale in nature (at the level of the Hamiltonian) due to scale invariance. We cannot have an infinite momentum as this will imply that some eigenvalues of observable properties would be infinite, which is not seen in practice. Therefore the Hamiltonian must be ultra-local in any finite and scale invariant field theory.

The second issue is related to Einstein's result that nothing can travel faster than the speed of light. If a Hamiltonian (which must be ultra-local in time since it is associated with the differential operator with respect to time) did contain fields separated by some distance but at the same time, then they would be communicating with each other instantaneously and information would proceed faster than light. This means that this situation is impossible. The fields have to be at the same place at each particular time. Thus the Hamiltonian has to be local. (This is, of course, related to the underlying Lorentz symmetry, which is the primary reason why nothing can travel faster than light.)

This, of course, means that interactions are simultaneous. An electron emits a photon and simultaneously we have a photon and an electron with a different momenta. It also implies causality. If we did have a term in the Lagrangian or Hamiltonian such as $\psi^\dagger(x)A_\mu(y)\psi(z)$, then we would have an electron decaying at a position x, a photon appearing at position y and an electron at position z. The electrons would appear to disappear into nothing or come from nothing. We would not even be able to relate for certain that the appearance of one electron was related to the disappearance of any particular electron elsewhere, or any electron. We would have no causality, or at best some Humean causality. But that is not the case; it is not what we observe and not what the theory (in particular the symmetries of the theory) demand. Every new particle (or particle anti-particle pair) originates from the decay of an already existing particle.

A further feature of quantum field theory is universality. One electron anywhere in the universe is the same as another electron somewhere else in the universe. This is keyed into the commutation (or anti-commutation) relations between the particle creation and annihilation operators. These commutation relations distinguish between particles of the same kind and particles of different kinds[6]. This, of course, means that there are 'kinds' in nature, and that they are not merely matters of our own construction or convenience, since the commutation relations have real physical consequences such as the exclusion principle. Similarly, one photon is the same as any other photon; the gaps between energy bands in a solid are the same whichever solid of the same kind we are discussing, and therefore the colour yellow in one banana is the same as the colour yellow in another banana, since the aspect of colour that is resident in materials is derived from the gaps between the bands of energy eigenstates of those materials. This rules out nominalist or conceptionalist views of reality, at least at the level of the physical world: realism is the only possibility consistent with quantum physics.

This is, of course, very different from the picture of classical physics, which is consistent with nominalism, and in particular its mechanistic interpretation, which some have believed to imply nominalism since in classical physics each individual particle is unrelated to the others and similarities of properties can be thought to be mere coincidence. Quantum physics rules this interpretation out (although we need different arguments against a nominalist interpretation of non-physical entities, such as biological species or ethics; although I firmly believe that even in cases such as this realism is the only viable option. Biological species I have already discussed; I will discuss ethics in a later chapter).

Another consequence of the universality of the electron field is the property of cluster decomposition: that is that the coupling between the electron and photon field (as an example; this may be regarded as being related to the probability that an electron emits

[6]The distinction between *kinds* I refer to here is the distinction between electrons, muons, red up-quarks, photons, and so on.

a photon within a certain time interval) is independent of what else is happening in the diagram. Equally, each propagator is the same function of spinor index and momentum for each type of particle whatever else is happening in the diagram and wherever it is in the diagram. Each particle is independent of what is happening around it. This stands against many metaphysical theories which depend on the inter-connectedness of all things (prevalent, for example, in those philosophies bastardised from Eastern religions): connections between different particles of matter are only through local interactions.

The next key finding of quantum field theory is that particles may be created and annihilated; that particle-anti particle pairs may be arise from the decay of other particles. Each individual lump of matter is not indestructible; the decay products need not be parts of the original object (as we would imagine decay in a classical picture). When a photon decays into an electron/positron pair, it is not that the electron and positron were somehow physically part of the original photon, nor do they come from nothing; rather the best we can do is say that they existed in the photon potentially, as a intermediate state between actual existence and actual non-existence. This non-indestructibility of material particles is implied by special relativity, and a key element for quantum field theory. Relativistic quantum mechanics, which does not allow for this phenomena, is mathematically inconsistent.

Another important aspect of quantum field theory is symmetry. Symmetry is key in modern physics. It is the key principle used to construct the Lagrangian, on which everything else depends. Noether's theorem relates symmetries to conservation laws; the breaking of symmetry (i.e. symmetries that are almost, but not quite valid) allow us to extract physical observables such as the masses of composite particles from the theory, spontaneously broken symmetries (symmetries broken by the choice of minimum energy state) also lead to important consequences. The symmetries can be both global (applied to all points in space/time simultaneously, and (more importantly, and more mysteriously) local (applied to each individual point in space time as much as possible independently of what is happening elsewhere). The symmetries of the standard model Lagrangian are translation (add a quantity to all the particles in the universe), rotation (around any point in space time, this includes Lorentz transformations – changes in velocity – which can be though of as rotations between space and time[7]), scale invariance (multiplying every position by a constant, no matter where the origin of the coordinates is), gauge invariance (unitary transformations of the matter fields), and reflection (again, with the origin at any position, simultaneously setting $x \to -x$, $t \to -t$ and converting particles into antiparticles). These are all present at the level of the Lagrangian, although some are broken once we introduce particular configurations of matter field (before we reach thermodynamic equilibrium). There are also various other approximate or broken symmetries which are also important in fundamental physics. These symmetries are something used to build physical theories; the correct symmetries can be identified by physics, but physics cannot explain why these symmetries are present. This should be a task of metaphysics, and a good test of whether a metaphysical theory is correct is if it can account for these symmetries (a necessary but not sufficient condition on any metaphysics).

Scholastic metaphysics was abandoned in the sixteenth and seventeenth centuries because it compared unfavourably with the mechanistic world view, which seemed to be more successful in spawning a physics which a agreed with observation. Most systems of philosophy developed since then have used mechanism as a starting point, either assuming some of its tenets, or reacting against it in some way. Throughout this period, it was assumed that scholastic metaphysics had been questioned by the early mechanical philosophers and found wanting, and thus it has not been considered as viable. Formal and final causality in particular have been abandoned by philosophers (except for those dinosaurs who refused to give up on scholasticism during the enlightenment, despite the evidence against it).

[7]Albeit with an imaginary rotation angle.

11.3. Mechanism and Quantum Field Theory

Now we are in the position to compare mechanism against the best physics of our era, quantum field theory. Of course, we do not have the full theory. It may be that quantum field theory is wrong. But given our current knowledge, and its enormous success, that seems to be unlikely. Most extensions to modern physics simply apply field theory to more situations: different types of particles, or strings in multiple dimensions rather than point-like particles in four dimensions, and so on. Should quantum field theory be overthrown in the future, then my own conclusions would have to be questioned, but for now we have to proceed with the best we have available.

In an earlier chapter, I defined a mechanistic world view as one which accepted the majority of the following axioms:

(1) That all compound objects can be reduced to the sum of their parts.
(2) That the effects of these parts can be determined solely by their extension, their motion, their duration, and perhaps a limited number of other qualities.
(3) That matter is fundamentally indestructible and cannot be created or destroyed.
(4) That the only motion of significance is locomotion, the movement from one place to another.
(5) That the motions of matter can be explained according to external fixed laws which operate universally.
(6) That these laws are deterministic.
(7) That space and time are independent of matter and can be mapped to a geometrical representation.

The first of these axioms is false, as we have seen. The state of a physical object can be described in terms of its energy levels, and the Hamiltonian of a compound object is not just the sum of the Hamiltonians of its constituents. The second axiom is also false: matter can be evaluated, and is most easily evaluated, in terms of energy and momentum eigenstates. This makes no mention of extension and duration of corpuscles, and locomotion only indirectly through momentum. The third axiom has been disproved by countless experiments. The fourth axiom is also false (given that motion refers to the scholastic definition meaning any change of state); changes in the system induced by the spontaneous creation and annihilation of particles or from one eigenstate to another are very relevant to physics, and more so in many ways than locomotion. The fifth axiom is also false, at least in the way it is usually interpreted in classical physics; whatever laws of physics there are are very different than the form (i.e. differential equations, such as Newton's second law) that the mechanists drew from this idea. The sixth axiom is false, since quantum field theory relies on the assumption that the universe is wholly unpredictable. The seventh axiom is usually assumed in interpretations of quantum field theory (including my own), although in practice measurement of position is characterised by the application of an operator to a Fock state which is dependent on the matter in the universe. However, given that we can apply the operator to any Fock state, and the operator itself is independent of the matter but does depend on an underlying geometry, it is best to assume that space and time should be thought of geometrically rather than materially.

In other words, mechanism as a philosophy is wrong. Only one of the seven axioms survives comparison with contemporary physics. Mechanism, and everything built on it, should be abandoned. Many of these axioms are wholly false. Some are maintained by some people. Probably the best known attempt to reconcile quantum field theory with one aspect of mechanism is the Everett or many worlds interpretation, which maintains a deterministic universe. This states that with every point where the universe acts unpredictability, it splits into two branches, one of which had one outcome, and the other the other outcome. We, of course, are conscious of only that branch of the universe we exist in; however there are parallel versions of us in different branches of the multi-verse which had a different history (as well as an infinite number of universes where we do not exist at all). This interprets the

probabilistic nature of quantum mechanics and quantum field theory from the perspective of frequentest probability. The system is deterministic as a whole because the evolution of the whole multi-verse is deterministic. We, however, see things as indeterminate because we only view one branch of the multi-verse.

Personally, I find this understanding implausible. It is not a scientific theory, because the alternative branches of the universe are not observable and thus the idea is not even in principle falsifiable. However, it lacks plausibility because there is no obvious reason why reality should branch in this way: it is not a natural outcome of any set of coherent meta-physical assumptions, and brings in various difficulties with our conception of space and time as being a coherent whole. The Everett interpretation attempts to resolve the measurement problem: the dichotomy between deterministic wavefunction evolution (fully understood by a differential equation dependent only on the physical states of material particles) and the indeterminate 'measurement' process (which is not described by a physical equation and cannot be determined solely by the physical state of the universe)[8]. The question is: Why is the Schrödinger equation insufficient to fully describe quantum mechanics? Why do we need this extra step of indeterminate (i.e. without a sufficient physical cause) wavefunction collapse? The Everett interpretation attempts to resolve this problem by saying that there is no indeterminate collapse of the wavefunction. Every possibility is sampled in one branch or another. The problem is that the Schrödinger equation remains insufficient to describe the physics. We still need to add something extra, lacking a sufficient physical cause, namely the branching mechanism. There is still indeterminism in the sense that it is just luck of the draw which of the branched detectors will get which result. Thus the Everett interpretation might solve wavefunction collapse, but it doesn't answer the fundamental philosophical problem. It doesn't remove the question, but only modifies some of the details.

I also find the Everett interpretation suspect because of its reliance on a frequentest understanding of probability, an interpretation that has numerous difficulties. If we have a choice between the assumption that reality is (as we observe it) unique but indeterminate, or determinate but not unique, then to accept that the universe is indeterminate is far less of a blow to our natural intuitions. The Everett multi-verse is too high a price to pay to maintain determinism.

I have also mentioned the trend in some modern philosophy to identify the essence of a thing with its observable properties. In classical mechanics, this sort of makes sense. In quantum mechanical language, all the observable operations commute, so any physical object would be in a simultaneous eigenstate of all observables, and must be in such an eigenstate. This eigenstate can be identified simply by listing out the observables, or properties of the object. Of course, the properties will still be frame-dependent. A property is a relationship between an observer and the object being observed. We cannot say that 'A London bus is red', because colour is not a Lorentz invariant: we always have to give the reference frame, so the correct statement is 'A London bus in its own rest frame is red'. In some other reference frame, the bus will appear blue, because the light will be blue-shifted for an observer approaching the bus at the appropriate velocity. Giving the inertial frame is the observer dependent part of the property; however these differences are linked by the mathematical transformations relating the different reference frames; so with knowledge of the properties of the object in one inertial frame (say the objects rest frame) we can compute what every other observer in every other reference frame perceives of the object. A complete description of any observation includes the observer's reference frame, and once this is done every observer's measurements are equivalent. Thus in classical mechanics, the essence of the object and its properties in a given reference frame are in a two way relationship with

[8]I put the word measurement in quotation marks, because strictly speaking the process occurs when a quantum state interacts with any macroscopic system, regardless of whether or not it is a detector. Decoherence forces the quantum state to adopt one particular basis. This leads to an indeterminate jump into one of the states in that basis. Every measurement will trigger wavefunction collapse; but collapse doesn't have to be triggered by some scientist observing something.

each other. I believe that it is still incorrect to identify them even here, but there is no observable consequence if this is done (we just introduce metaphysical errors). A complete list of properties is formally equivalent to a complete description of the object's form.

However, in quantum physics (both field theory and quantum mechanics), this approach is utterly wrong. Classical properties are related to the eigenvalues of various operators. These eigenvalues are not even enough to specify the observable operator itself - we need the eigenvectors as well – so the classical properties are not even a complete description of the being. But the key point is that in quantum physics, the object is either described by the wavefunction (in quantum mechanics) or the field (or, equivalently, the various creation operators that can be used to define the field), and these are in general wholly independent of the observation operators. What a wavefunction (or field) allows us to do is compute the probability that an observer in a specified reference frame will measure a particular property. The properties are not determined, but only probabilistic. This does not mean that the quantum particle is something weird, or something that it is not there unless we measure it. It merely means that it is different from a classical particle. It cannot be identified with its observable properties, even though the probability that we will observe a certain property can be derived from it.

11.4. Scholastic Metaphysics and Quantum Field Theory

So what of scholasticism? Scholastic metaphysics can be defined as one that incorporates the idea of potentia and actuality, admits to the four causes, and accepts the principle of causality and the principle of proportionate causality.

We have identified the different energy eigenstates of a system with the scholastic potentia. This identification is not new to me; the earliest person to do so that I am aware of was Heisenberg[9],

> These results seem at first sight to lead away from the idea of the unity of matter, since the number of fundamental units of matter seems to have again increased to values comparable to the number of different chemical elements. But this would not be a proper interpretation. The experiments have at the same time shown that the particles can be created from other particles or simply from the kinetic energy of such particles, and they can again disintegrate into other particles. Actually the experiments have shown the complete mutability of matter. All the elementary particles can, at sufficiently high energies, be transmuted into other particles, or they can simply be created from kinetic energy and can be annihilated into energy, for instance into radiation. Therefore, we have here actually the final proof for the unity of matter. All the elementary particles are made of the same substance, which we may call energy or universal matter; they are just different forms in which matter can appear.
>
> If we compare this situation with the Aristotelian concepts of matter and form, we can say that the matter of Aristotle, which is mere 'potentia,' should be compared to our concept of energy, which gets into 'actuality' by means of the form, when the elementary particle is created.
>
> ...
>
> In a similar way in quantum theory all the classical concepts are, when applied to the atom, just as well and just as little defined as the 'temperature of the atom'; they are correlated with statistical expectations; only in rare cases may the expectation become the equivalent of

[9]Though Heisenberg was not the last, see also (Smith, 1950; Hoenen, 1956; Smith, 2005; Wallace, 1996) for a sample of authors who have considered an Aristotelian understanding of quantum physics. However, these authors only consider quantum mechanics rather than field theory.

> certainty. Again, as in classical thermodynamics, it is difficult to call the expectation objective. One might perhaps call it an objective tendency or possibility, a 'potentia' in the sense of Aristotelian philosophy. In fact, I believe that the language actually used by physicists when they speak about atomic events produces in their minds similar 'notions as the concept potentia.' So the physicists have gradually become accustomed to considering the electronic orbits, etc., not as reality but rather as a kind of 'potentia.'
>
> ...
>
> One sees at once that this use of the word 'state,' especially the term 'coexistent state,' is so different from the usual materialistic ontology that one may doubt whether one is using a convenient terminology. On the other hand, if one considers the word 'state' as describing some potentiality rather than a reality – one may even simply 'replace the term 'state' by the term 'potentiality' – then the concept of coexistent potentialities' is quite plausible, since one potentiality may involve or overlap other potentialities [10].

Heisenberg wrote this in 1958, before quantum field theory was fully developed and established, but after the fundamental ideas had been proposed and the first details worked out. The late 1940s and 1950s were a time one of great excitement and progress in particle physics, both experimentally and theoretically, and quantum field theory was arguably the most important theoretical development of that era. However, the ideas were in their early stages at that time, and the process of development and understanding would continue into the late 1970s. Thus when he wrote this, enough of quantum field theory was known for Heisenberg's work to remain useful. Heisenberg himself was obviously one of the great physicists who inspired the emergence of quantum mechanics; and his was probably the last generation where the great physicists were also competent philosophers – knowledgeable about the debates in the philosophy departments. The generation that came to prominence after the second world war, which gave us field theory in its mature form, were not so philosophically aware. This was perhaps not surprising: quantum mechanics seemed to indicate the triumph of the experimental method over the philosophical method, and the 1930s, when this generation was trained, was a period of increasing specialisation in academia. Philosophy had failed, just stop worrying about it and calculate. The best physicists of the time were no longer competent in philosophy; the best regarded philosophers knew little of modern physics (although, obviously, there are exceptions, and a few who are expert in both philosophy and modern physics; these were, however, a significant minority, and not so influential in either field). I should also note that while Heisenberg suggested that this element of Aristotle's philosophical framework might be useful, he does not take the idea any further than equating energy eigenstates with potentia. I don't agree with everything in the above quote, such as his equating energy with prime matter.

In terms of fundamental particles, the notion of potentia also fits. It is difficult to imagine the process of photon emission in terms of mechanical categories. However, if we think in Aristotelian terms, it is straight-forward to suggest that an actual electron of a given momenta is potentially a photon and an electron of a different momenta; an actual photon is potentially an electron/positron pair, and so on. Change in quantum field theory can be viewed as natural if we use the same language that Aristotle made central to his thinking about of what change must involve all those millennia ago.

If the idea of potentia fits well with quantum field theory, then what of the other premises of scholastic metaphysics? Let me return to the list discussed in chapter 4.

(1) **There are sets of types of material objects.** Yes. Fundamental particles are indistinguishable from each other. We then can and should class together electrons;

[10] (Heisenberg, 1958), chapter 9; see (Silva, 2013) for a fuller discussion of Heisenberg's views.

protons; ethanol molecules and so on because their structure of energy eigenstates are not merely similar but identical to each other.

(2) **These types are not a human construction.** Yes. As discussed above, this is shown by the Pauli exclusion principle.

(3) **There is a unifying principle, form.** Yes. I have identified the form with the set of eigenstates of continuously connected Hamiltonians. It may be, of course, that there is more than just this involved in the form, perhaps describing some non-physical features, but I will leave that for the philosophers to chew over. The energy eigenstates describe all the relevant physical information, so as a physicist I am not so much concerned about whether there is any additional unphysical information.

(4) **There is a distinguishing principle, matter.** Yes. Physical objects are more than just a mathematical representation.

(5) **A material substance is the union of form and matter.** Yes.

(6) **A material substance is not simply the sum of its parts.** Yes. Hamiltonians of compound objects are not the sum of the Hamiltonians of its constituents. The transformations that lead to the appropriate effective field theory are not linear.

(7) **We can grasp and understand form in the intellect.** Yes. We can represent the form mathematically. This has been done for the simplest objects; it can in principle be done for any object.

(8) **We cannot grasp matter intellectually.** We cannot. We can describe how an electron behaves, enumerate its properties, but these are related to the form. The matter that individuates each individual electron cannot be expressed abstractly, so therefore it is not an object of our understanding.

(9) **Each form consists of many potentia.** Yes. The potentia have been identified with the momentum eigenstates of the Hamiltonian.

(10) **Change (or motion) is defined as the movement of matter from one potentia to another.** Yes. For example, the emission of light or heat is caused when an electron moves from one eigenstate to another, releasing a photon.

(11) **There is an actual state.** Yes, but with complications caused by the possibility of linear superpositions. Before renormalization, we can define states, we have to select the representation of the basis, and this representation is not unique; consequently objects might be in a superposition of eigenstates. However, renormalization fixes one particular basis as physical and restricts this freedom. Physical beings are exact eigenstates in one particular basis; if this basis coincides with the renormalized basis, then only one state is actual at a time. If not, then we have to add to Aristotle the possibility that the being is in a superposition of potentia.

(12) **Another possible type of change is the creation or destruction of material substances.** Yes. Seen every day in particle accelerators.

(13) **There is an efficient cause for every change of state.** Yes. Guaranteed by the conservation of four-momentum. Each particle has its own efficient cause, since each particle is emitted from some other particle or emerges from the decay of some other particle, or the mutual annihilation of two (or more) particles (there is a four gluon interaction allowed in the Hamiltonian for the strong nuclear force, so in principle three gluons can merge into a single gluon). In each case, we can identify the efficient cause of a particle clearly. Each connected Feynman diagram represents one possible physical process. One of those processes happens in reality, and that process gives us the efficient causes of the final state. We cannot, of course, identify which process happened; but this does not mean that there was not some process which happened. There is always an efficient cause, even if our knowledge of the chain of causality is only probabilistic rather than with certainty. The principle of causality, that only something actual can induce

the change from potentiality to actuality can be mapped to the idea that a change in energy states requires that there is some initial state that needs to be excited or decayed, as guaranteed by the conservation of four momentum and the absence of any wholly disconnected diagrams in the perturbative expansion (which would involve particles emerging from the vacuum if it happened). Of course, in reality the QCD vacuum state in particular is not empty but contains various gluonic structures; however these structures are not nothing, so even if they induce some creation or annihilation of particles, those particles are not coming from nothing but something present within the vacuum.

(14) **A potentia has a limited number of final causes.** Yes. Enumerated by physicists as the decay channels of a particle. These allowed and forbidden interactions are determined by the conservation of momenta and electric charge, and summarised in the Feynman rules. The scholastic notion of final causality is just a list of possible outcomes. The physical concept of decay channels goes beyond this, by identifying the various cross sections, or probabilities for each possible outcome, which can be computed from knowledge of the theory and the fundamental physical parameters. Thus the scholastic concept is more general than the concept that arises from physics. The idea of certain decay channels implies final causality; final causality does not imply the concept of decay channels. Thus in this respect the modern physicist has done better than the medieval philosopher. But the medieval concept is certainly present in modern physics. The principle of proportionate causality, that every effect is contained potentially or virtually within its cause, can be mapped to the idea that only certain decays are possible, so every photon can be said to contain an electron positron pair virtually following the scholastic terminology without making any compromises to the physics. Equally important is that no two particles can decay into nothing, as guaranteed by the conservation of momentum.

(15) **Properties can be understood from knowledge the form.** Yes, but with complications. The properties can be and are computed from knowledge of the energy band structure of a material. The complication is that sometimes these properties are undetermined, if the potentia is a linear superposition of the eigenstates of the operator representing the observable property. In this case, we cannot determine precisely the property, but only say that there is a certain probability (given that it is in that potentia) that we will measure one value, or a certain probability that we will measure another. However, over a large enough sample of objects, the spectrum of this property will fall in a determined frequency distribution. If the potentia is in an eigenstate of the object, then the properties are determined from knowledge of an object's state.

(16) **Not every property is essential to the form.** Also correct. For example, potentia in different eigenstates of the Hamiltonian have different energy.

In short, while mechanism failed in almost all of its principles, Aristotle comes very close to succeeding in everything, with only some minor complications arising from the comparison. The complications arise from the superposition principle, about which Aristotle had no real idea. A quantum state can have amplitudes in more than one state at a given time, something which seems unintuitive. There are several ways of avoiding this. One is to say that the quantum mechanical state merely references our knowledge of the world, which leaves open the idea that the real physical world is in a single specified state, but our knowledge of it is imprecise and this imprecision is given a precise representation by writing quantum mechanical states as superpositions. However, even if the Fock state merely describes the state of our knowledge, that does not mean that the real world is necessarily free of superpositions. Secondly, we could extend the concept of potentia to not only include every energy eigenstate, but also every possible superposition of energy eigenstates. Thirdly, it

may be that our intuition that a particle may be in only one state at a time is wrong. There may be more possibilities than this. The first two possibilities are certainly consistent with scholastic metaphysics, so linear superposition is not a great hurdle for the scholastics, and they should be able to incorporate it into their philosophy with just a minor modification.

It should be noted however, that scholastic causality is just about the only (viable) notion of causality that survives quantum field theory. In particular, the distinction between substance and event causality (see section 4.5) is crucial. The versions of causality that depend on events, sufficiently popular in enlightenment physics that they are sometimes today presumed to be the only versions of causality, are no longer valid because the events, namely the decays of particles, are unpredictable. We cannot say that there is a physical or material cause for any particular event (there may, of course, still be an immaterial cause). Hume's confusion between causation and correlation, the alternative to event causality that emerged from the enlightenment, is not forbidden specifically by quantum field theory, since it is only probabilistic, but, with its denial of theoretical insight, is not viable in either classical nor quantum physics. It remains invalid for the same reasons that it was a non-starter in classical physics.

It is wrong, however, to conclude from this that all forms of causality are denied by the quantum theory. Substance causality, including Aristotle's efficient causality, is still valid. The principles of the locality of physical interactions (substance-locality, i.e. the locality of the Hamiltonian operator and the acts of creation and annihilation it describes) and the conservation of momentum guarantee that every physical state emerges from another physical state, and that the physical causes are sufficient to explain the effect. Events, on the other hand, such as when a decay occurs or which state is selected during a measurement, cannot be explained in terms of physical causes, and two separated events can also be correlated in a such a way that violates event-locality (see the discussion of Bell's theorem in section 15.5). But that's fine for Aristotle, since his philosophy only requires substance causality.

Earlier, when discussing the beginnings of geometrical physics in the fourteenth century, I noted that expressing physics in terms of geometry reduced the scope of Aristotle's physics, and in particular cut out notions of form and causality, not because they were disproved but simply because they could not be expressed in terms of the mathematical structure that was used to describe locomotion and dynamics. From the success of that physical approach, it was assumed that all of physics could be reduced to extension and motion, and perhaps a few other quantities such as mass which could also be assigned numerical values. Now we find that we have expressed causality, including formal and final causality, mathematically. I still believe that causality cannot be treated symbolically using the mathematics available to Bradwardine and Buridan or even Newton, and thus these scientists were right to seek an alternative approach. But later advances in mathematics changed things: the advances, starting from the nineteenth century, in group and set theory, topology, probability and operator theory did provide a framework in which we might hope that the rest of Aristotle's metaphysics can be expressed mathematically. I have just scratched the surface of that project here; the idea for constructing time evolution operator from creation and annihilation operators gives us the framework to express final and formal causality mathematically; searching for the eigenstates of the time evolution operator, that is meta-stable states of the system, allows us to understand the set of all such states as form, and the process of actualisation in terms of transitions between those states. Because we can create effective field theories, combining the creation and annihilation operators of the fundamental particles of matter into creation operators for increasingly complex composite structures, using these mathematical and statistical techniques to trace out their distinct energy bands, giving another set of potentia and formal and final causes for atoms, molecules and solid crystals, bringing all of physics and chemistry into the same framework. These energy bands can be used to compute the properties of any material substance from knowledge of the form (or the

possible potentia, or the eigenstates of the composite Hamiltonian). Using the mathematics of topology, we can distinguish between continuous and discontinuous deformations of the Hamiltonian (and thus the potentia), and identify the various different kinds or species of material beings. In short, I believe (though it will take far more work to establish this and work out the details) that what was not true in the days of Aquinas and Bradwardine is now true: it is possible to create a precise mathematical representation of not only the physics of motion, but also the entirety of metaphysics.

So what of the reasons why scholastic metaphysics was abandoned? The strongest of these was from the principle of inertia: that a state had to be defined in terms of both position and momentum, but that position continually changed without anything directly changing it. This is undermined in quantum mechanics and field theory by the observation that states can be defined in terms of momenta alone, and with the conservation of momenta the principle of causality remains. Other arguments based on differences between classical and Aristotelian physics are also no longer valid since classical physics is incorrect.

Other arguments were that the Aristotelian system was occult and ill-defined. Once combined with quantum field theory, though, the definitions (at least with regards to material substances) become solid. There were objections to the teleological aspect of the scholastic system, but some of these failed because they were based on a confusion between teleology and design, and others fail because they assumed that the system is deterministic and thus efficient and final causality could be identified with each other. The first of these classes of objections is just an error of logic, the second falls when we realise that physics is indeterminate. Objections to form rely on the idea that beings can only be in one defined state (and thus there are no potentia), but this objection also fails because its premise is disproved by quantum physics. Objections to the structure of causality fail because the alternative views of causality proposed by enlightenment philosophers, whether event causality or Humean causality, fall short of reality, while scholastic causality remains.

Others questioned the lack of progress of scholastic physics. This is true, but partly because when scholasticism was at its height the necessary mathematical tools had not been developed. Of course, the scholastic scholars were partially to blame for this: they (by giving too high a regard to Aristotle's views on the matter) did not adequately explore the idea that physics should be treated mathematically, nor attempt to develop the tools to do so. This, however, is the fault of the scholars more than their system. Secondly, there were mistakes in the medieval scholastic system, most notably its view of space and time, which hindered the progress of the scholars, but these can be corrected without the wholesale abandonment of the metaphysics. Just as that the ideas behind classical mechanics led to rapid progress in physics did not mean that those ideas were right, that the ideas behind scholasticism did not lead to rapid progress at that time did not mean that those ideas were wrong. Equally, that final causality led people down blind alleys did not mean that the idea of final causality is wrong, only the uses to which it was put in medieval times.

In short, the objections raised against scholastic metaphysics at the emergence of the modern period fail: either they were not rigorous arguments in the first place, or they were based on ideas in classical physics which have subsequently been shown to be false. Indeed, it is interesting that classical physics fails in precisely those aspects which were used as evidence against the scholastic system.

11.5. Summary

Mechanism is inconsistent with quantum field theory. Kant's approach, which is based on the premise that objects can be defined by listing their properties, is disproved because it cannot cope with indeterminate properties. The empiricists are disproved because their fundamental premise that our understanding is limited to only that which we can observe or simple geometrical ideas takes no account of our success in understanding cause, form, property and finality in ways which the empiricist claims can't in principle be done. The

philosophical relativists are disproved because quantum field theory is objective; it does not depend on the social status or preconceived prejudices of the observer. The existentialists are near enough disproved because they cannot explain why our experiences obey these beautiful but profound laws. The Toaists are disproved because their four element and five tastes theory of matter is inconsistent with what physics has found. But Aristotle, even though he wrote over two thousand years before quantum field theory was developed, came very close to getting it right. Aristotle alone stands and survives. Indeed, in his philosophy quantum field theory makes sense (or at least more sense than it does in the more modern philosophies). The reason for this is obvious; Aristotle based his arguments on the fundamental principles of motion and change, of similarity and difference, which every study of physics must presuppose. Other philosophers started from less certain first premises, and naturally enough went astray.

Of course, my selection of Aristotle is not to idolise the man. There were numerous ancient philosophers, both in Greece and other countries, developing numerous different schemes (some of which resembled the enlightenment philosophies), between them covering most of the different possibilities. With such a large sample to choose from, it is not so unlikely that one of them could serve as a basis for a philosophy of quantum field theory. What I have done is select the one from that sample which best serves my purposes. Of course, Aristotle did not reach his answer purely by luck: he was a meticulous and careful thinker; and that served him in good stead.

But does that mean that we should accept the traditional view of metaphysics without criticism? No. I am not advocating that. If we accept it, it should be on its own terms, because its assumptions are reasonable and it passes the test of consistency with physics, and not on account of any traditional authority. Certainly we should re-examine all the details of the old system, and should be quick to point out where the medieval scholars did get things wrong. Obviously, replacing Aristotelian physics with quantum physics is not a small change, with ramifications throughout the system, even if it the broad details of the metaphysics are consistent with modern physics. We should not accept the medieval writings as they stand: they must be re-thought from first principles in terms of the new physics. There are other differences as well; I have already discussed the conception of space and time and superposition, but another place where the scholastics were wrong (or at least imprecise) was in their understanding of the theory of knowledge.

One of the big questions raised by the enlightenment philosophers is to ask what the relationship between our knowledge and the real world. Aristotle had assumed that by grasping and understanding form we could understand everything of importance about reality, in other words that it is possible to fully understand reality. Kant is the best known member of the movement that questioned this, positing a clear distinction between the world that is and the world of our knowledge, using examples where we are mistaken to confirm that the two should not be identified.

I believe that Kant was right to question how much what we know corresponds to reality, but wrong about how he interpreted this, and consequently almost everything else. He made two important mistakes. The first was to deny the existence of potentia and consequently form; he implicitly accepted that knowledge and belief had to be perfectly precise. The idea that we could and ought to say that the real world is one of a given set of possibilities was not discussed in his work; thus we can be accurate but not precise. For example, if we see a man in a distance, we ought not (if we are thinking clearly) say 'Dark hair, fair complexion, good build, reasonably tall, brisk walk - it must be Derek', only to discover later that it was not him, but instead we can legitimately say 'Dark hair, fair complexion, good build, reasonably tall, brisk walk - it must be Derek or somebody who resembles him'. That statement remains true even when not-Derek approaches closer – it was somebody who looks like Derek. This is to introduce into our thinking the experimental error of the observation that all scientists are familiar with. The mistake in identification is not an error

in observation or imagination but of sloppy thinking, not properly identifying or dealing with uncertainty. The mistake that Kant made was that he took an all or nothing approach – either the world of knowledge was to be identified with the world of reality, or the world of knowledge was separate from it. Kant choose to keep it separate, which, of course, lead to the philosophical problem of knowledge – how does what we know relate to what is. However there is an alternative, which is to use a Bayesian probabilistic model. Instead of saying 'The world is like this', we say that 'Given these observations, there is a certain probability that the world is like this', with the probability used as a way to classify our degree of uncertainty. Each new observation modifies the initial information from which we compute the probabilities and thus the probabilities themselves. Rather than being separate from reality, we should (if we are thinking straight) classify it as best as we can probabilistically. This avoids Kant's error of disconnecting our knowledge entirely from reality and also the naive error of identifying the two.

At least that would be the best approach if we were infallible but not omniscient beings: limited in knowledge and awareness of facts, but not making mistakes in the interpretation of those facts. Of course, in practice we do make mistakes, either in logic, method, incorrect premises or our observations. Here we just have to be humble, aware of our fallibility, and try our utmost to allow and correct for it as much as possible. Of course, even here we can weight different beliefs differently: we can have high confidence in a simple calculation which we have long toiled over, checked numerous times and where there is no possibility of hidden assumptions hiding additional complexities we have not accounted for and which agrees with what other people have concluded and with experiment; and a lower confidence in a complex calculation which we rushed through and nobody else has attempted and has not yet been tested against experiment; and no confidence in something that disagrees with experiment or makes bad predictions.

So what about our knowledge of physics? There are two aspects of this. The first is our knowledge of the Hamiltonian operator, Feynman rules, and thus the formal and final causality. This I believe that we can in principle know fully, though in practice we don't for three reasons. Firstly, we cannot be certain that we have the correct mathematical form of the Hamiltonian operator; there could well be some new physics either in the form of a new particle or some gravitational effect which will come in and bite us on the backside. My belief is that one day we will uncover the theory of everything, and this will not be an issue. Although how can we be certain that we have indeed uncovered the theory of everything? This is something which is less clear. It may be that the mathematical form of it is such that we will realise that it has to be that way, and slap ourselves for our stupidity in not seeing it earlier; more likely though there will always be an element of doubt that there could be a new particle just beyond the sensitivity of our best experiments. Obviously my belief that the theory can and thus will be uncovered is based on nothing but blind optimism and could be mistaken. But if it merely could be mistaken, then the exactness of our understanding can be treated probabilistically: assuming that we have correctly understood quantum theory and the underlying symmetries of nature, we can put bounds on what what new physics could be, and use these to restrict the range of possible influences it has on our calculations.

The second difficulty is that the Hamiltonian, and thus formal and final causality, depends on various parameters, and these cannot be computed but only measured experimentally, and each measurement contains an experimental error. Once again, however, we can treat these errors in a probabilistic fashion, using the range and distribution of possible values of these constants to compute a probabilistic distribution of the energy eigenstates and thus the formal and final causes of nature.

The third difficulty is the sheer challenge of some of the calculations, which we are not likely to be able to perform exactly, but only up to a certain precision, and this will also introduce uncertainties into our final result. Thus our knowledge of form and finality can in principle be exact but in practice will only ever be imprecise, although that imprecision

can be treated mathematically with precision by expressing it as a probability distribution (probability being the art of being precise about imprecision).

Thus our knowledge of physical law will always be imprecise in at least some details. But imprecise is not the same as being inaccurate. The imprecision is bounded so that reality lies within a known range. The imprecision will decrease over time as better measurements are performed and the next layer of the theoretical approximation is unravelled. And, there are aspects of the law, such as the overall framework and the symmetries that define it, which will be known both precisely and accurately. It is these aspects which provide the link with philosophy and from that other fields of knowledge.

The second aspect of our knowledge is efficient and material causality. What is the state of the universe right now (in a given reference frame)? What processes got it here? Or, alternatively, given the state of the universe now what is it going to be at a given moment in the future? Questions such as this cannot even in principle be known exactly (except perhaps by God), because of the indeterminacy of physics. Here we can only ever express the options with given probabilities; the same sources of error that relate to formal and final causes arise here, but also the more fundamental barrier of unpredictability. This is why in quantum mechanics and field theory the key element is only a probability amplitude.

Thus our knowledge of reality is only probabilistic. This does not mean that reality itself is probabilistic or random in some way -- on the basis of our discussion so far it could be, but is not necessarily so. My own belief is that nature in any given time in a given reference frame is in a single well defined state, and it is only our knowledge of that state which is uncertain. This limitation of our knowledge is something that neither Aristotle nor the scholastics thought through rigorously, and any new approach to philosophy would have to do so.

I do not want in all this to give the impression that scholastic metaphysics is derived from physics. It is more fundamental than physics; metaphysics comes first in the order of being, and the assumptions behind physics are derived from it. Instead, the study of physics is necessary as a check and confirmation of the metaphysics, rather than a justification of it. If the metaphysics is true, then the physics should be consistent with it. But the metaphysics is more fundamental, and derived from various basic assumptions, meaning that the metaphysical system does not only apply to the material world, but also anything immaterial as well (if there is anything immaterial).

To distinguish between various metaphysical systems, it is useful to list four basic assumptions. My discussion here is obviously a little simplistic, however I believe that it still is correct in broad, if not fine, detail.

(1) Reality is wholly rational.
(2) Reality is wholly knowable.
(3) Substantial change is possible.
(4) Space and time should be understood geometrically.

Aristotle's metaphysics is a natural outworking of the first three of these assumptions (with 'substantial change' defined following Aristotle's terminology). Aristotle himself rejected the fourth assumption it; his own physics is obviously non-geometrical; however Aristotle's metaphysics is blind to this assumption, and does not depend on its truth or falsity. If you accept the first three assumptions then you will finish with something close to Aristotle's system; at least in broad outline. There are many finer details which can still be discussed, but I do not think that one can avoid the basic ideas of actuality, potentiality, form and finality.

There are three main alternatives to Aristotle's metaphysics which have been proposed (that would accept the existence of reality; denial of reality is, I think, wholly implausible): Hume's system, Kant's system, and the mechanistic world-view which inspired (though not in full agreement with) the writings of Descartes.

Hume, as I read him, denies the first premise of these four. That is not to say that he regards the universe as irrational, merely that he does not regard the rationality of the universe, nor the particular way it is outworked, as something that can be demonstrated, so we need to develop a system that is general enough to cope with a partially irrational universe. The success of physics, however, undermines his whole system.

Kant agrees with the first and fourth premises, but disagrees with the second and third. He denies the possibility of substantial change, and denies that reality is knowable. Once again, the success of physics makes Kant's system untenable: something about reality certainly is knowable.

The mechanists accept the first, second and fourth premises, but deny the third, saying that substantial change is impossible. Unlike Hume and Kant, I believe that mechanism is at least a rational possibility, in the sense that it is self consistent and consistent with the basic assumptions needed to perform scientific investigation. Were classical mechanics the correct description of physics, then there would be a very strong case for mechanism; I do not believe that scholasticism would be completely ruled out (classical physics is, after all, merely the $\hbar \to 0$ limit of quantum field theory, and quantum field theory at any $\hbar$ is consistent with the scholastic system), but, at least with respect to compatibility with physics, the mechanical world view would be preferred. Mechanism fails only when compared to experiment, and the results that reality is indeterminate, and that particles can be created and destroyed, different potentia exist, and thus substantial change is possible.

What the success of quantum field theory tells us is that reality is rational, that substantial change is possible, and that space and time should be understood geometrically. Acceptance of the first and third of these premises naturally leads to some variant of Aristotle's metaphysics; acceptance of the fourth premise means that the physics will develop differently from Aristotle's. Clearly to arrive at the standard model we need to add in further details and premises, but those are not so relevant at this level of the discussion. The second premise is a bit more interesting; and I believe that it should be replaced with

(2) Reality is knowable, but some aspects of it only probabilistically (with probability to be interpreted as a measure of uncertainty).

Aristotle did not sufficiently consider the difference between reality and our knowledge of it. By stating that our knowledge of reality is at best probabilistic we avoid the errors of both Descartes and Kant. Kant envisioned a complete separation between the world of our knowledge and the world of reality; this leads however to the problem of how we can trust our knowledge at all. By saying that we can understand the world probabilistically, we confirm the distinction between our knowledge of reality and reality itself, but also maintain the needed link between the two. Measurement and observation connect us with the real world, and adjust the probabilities by adding in new evidence as a prior.

Obviously the modification of the second premise and acceptance of the fourth means that we cannot accept the whole medieval construction; nor should we. The philosophers need to redevelop their ideas in line with these changes. However, the acceptance of the first and third premises means that much of the scholastic system and terminology still stands. Moreover, the first and third premises are general: they do not simply apply to physics, but instead must apply beyond physics as well. If substantial change is possible in physics, then it is possible in general. This means, of course, that the medieval arguments concerning God need to be considered. We have shown that their premises (baring some qualifications concerning our knowledge) are reasonable, and likely to be correct. The only question then is whether the logic and reasoning was a sound development of these premises.

One of the further differences between mechanism and Aristotelian philosophy is that in mechanism particles are blindly subject to laws external to them which describe their motion without reference to the internal nature of those particles, while the Aristotelian causes are either wholly internal to the particles (pure Aristotelian philosophy) or some compromise position (Thomism). The potentia of a substance are part of the nature of

that substance, and since one set of final causes describes possible relationships between the potentia, it seems that the dynamics of the theory must be part of the internal nature of that being: form and finality are linked. How a particle interacts with other substances is part of the definition of that particle, and thus its final causes are in some way intrinsic to it. So it is part of the defining nature of an electron that it is capable of emitting a photon and moving into a different energy level with a certain likelihood, rather than there being some external law or agent that wholly compels it to emit the photon. On the other hand, since the changes are indeterminate, they cannot be wholly internal to the particle. For example, we might imagine each particle having an internal 'clock', running down and the particle decays after it reaches zero: if this were true, then one might believe that the reason why the particle decays is wholly internal to its nature. But it is false, suggesting that a compromise position is better. Two possible compromise positions are that an external agent, acting consistently with the particle's intrinsic final cause, directly gives impetus to the particle, or that an external agent actively sustains the particle's existence and nature, with the final causes part of that nature, and the decays occur due to a change in the actions of that external sustaining agent. Of course, although I use the word 'agent' here, we should not bias ourselves by presupposing that it is some rational (or personal) agent; nor should we rule that out by presupposition. This will be discussed further later.

11.6. Principle of Causality

At the heart of Aristotelian metaphysics is the principle of causality, defined as

CONJECTURE 11.6.1. ***Principle of causality:*** *Every change in state requires an initial state that already exists.*

Proof:

The proof requires that a change in state is only possible if there is an initial particle which it acts on. The operator describing the change of state is constructed from creation and annihilation operators; the creation operators describe the final state after the interaction, and the annihilation operator describes the initial state before the interaction. If there is no such initial state, then that term in the interaction Hamiltonian has no effect. A change of state requires that we annihilate the initial state, and create the final state, with, perhaps, some additional particles. The principle of causality states, in effect, that there can be no term contributing to any operator describing a change of state constructed solely from creation operators (which would imply creation of new particles without requiring an initial state) or solely from annihilation operators (which would imply the destruction of particles into nothing). Of course, quantum field theory, and physical law in general, just describes causal interactions between physical beings. It may be that a supernatural being (if there is any such being) could act in ways other than that described by quantum field theory; but if so, that supernatural being would itself a) already exist and thus satisfy the requirements of the principle of causality, and b) act as a direct efficient cause of the final state. Thus, even though I concentrate on the physical case here, if there is more to the cosmos than just physics, that would not override the principle of causality.

We defined a state in terms of the eigenstates of the Hamiltonian operator, $\hat{H}$ with eigenvalues E_i, and if we take the time dependence out of the coefficients to make it explicit, we see that,

$$|\psi\rangle = \sum_i c_i(t_0) e^{i\frac{E_i(t-t_0)}{\hbar}} |E_i\rangle \,, \tag{379.1}$$

with

$$\hat{H}\,|E_i\rangle = E_i\,|E_i\rangle \tag{379.2}$$

Here the state $|\psi\rangle$ is a Fock state recording all the particles in nature.

For a system in a fixed eigenstate of the Hamiltonian, i.e. where one of the c_i is 1 and the others are zero, the state of the particle does not change in time, except for an unimportant shift in the phase which does not contribute to the probability. This means that if the system is in a given eigenstate of the Hamiltonian, then it will remain in that eigenstate permanently.

What if it is in a supposition of eigenstates, or the Hamiltonian operator itself changes? It will nonetheless be in the eigenstate of some particular operator, say $\hat{H}'$, since every state is the eigenstate of some operator. The system is a combination of particles; the particles are meta-stable, so the difference between $\hat{H}$ and $\hat{H}'$ cannot be so large that there would be constant change. The difference between the true Hamiltonian $\hat{H}$ and this approximate Hamiltonian $\hat{H}'$ will, like both those operators, be built as the sum of terms containing various creation and annihilation operators, and will contain terms such as $\Delta\hat{H} = \hat{a}_j^\dagger \hat{A}_{ij} \hat{a}_i$ for an operator $\hat{a}_i$ which represents the annihilation of Fermions and creation of anti-Fermions, an operator $\hat{a}_i^\dagger$ which represents the annihilation of anti-Fermions and creation of Fermions, and some other operator $\hat{A}_{ij}$ which might represent the creation or destruction of gauge Bosons or other Bosonic fields, or might just be some number (these will not be the particle/anti-particle creation and annihilation operators found in either $\hat{H}$ or $\hat{H}'$, but rotated into some different basis). i and j represent indices distinguishing the state of the particle; in this discussion I will restrict myself to just spin and either the location or momentum of the particle, but the reader can mentally add colour or other additional indices if they like to make the discussion more complete: it will not affect the argument. $\Delta\hat{H}$ drives the change in the system. Change is only possible if this operator acting on some initial state $|\psi\rangle$ gives something which is neither zero nor proportional to $|\psi\rangle$. Any term within this Hamiltonian operator containing an annihilation operator when acting on the empty state will give zero (and could not produce any change), but in principle, one can have a term which contains a Fermion creation operator, anti-Fermion creation operator, and perhaps a Bosonic creation operator.

Could this drive matter emerging without an initial state? The objection to this is built from four principles: firstly that energy and momentum are conserved, secondly that energy is positive[11], thirdly that the state with no particles has zero energy, and fourthly that every existent particle carries some energy. Put these together, and such creation from the vacuum is impossible.

That energy is always positive follows because, just as the momentum is the measure of how something moves through space, so energy is the measure of how something moves through time, as follows from its definition as the eigenvalue of the differential operator with respect to time. A positive energy eigenvalue refers to something moving forwards in time (from past to future); a negative energy would refer to something moving backwards in time (from future to past). But we know from experience, and from the definition of time as a measure of change, that objects only move forwards in time. Therefore energy has to be positive, and the positivity (or, in some possibilities inapplicable to the world we live in negativity) of energy, is a feature of any valid coordinate system, flat or curved.

Like both $\hat{H}$ and $\hat{H}'$, $\Delta\hat{H}$ must be local, i.e. all the operators refer to particles at the same location, and we must also, as in both H and $\hat{H}'$, integrate the operators over all of space and time. In quantum field theory, energy and momentum refer to the eigenvalues of the differential operators. So, when we write down the creation operators as a linear sum of momentum states,

$$\hat{a}_{\vec{x}} = \int \frac{d^3p}{(2\pi)^3\sqrt{E}} \hat{a}_{\vec{p}} e^{\frac{i}{\hbar}\vec{x}\cdot\vec{p}} \tag{380.1}$$

[11] In this section, the word *positive* is exclusive of zero: all existent particles carry energy.

Before performing any calculation, we need to fix to some gauge (specify the phase degree of freedom of our creation/annihilation operators), and it is convenient to choose a gauge where the time component of the electromagnetic potential, $\hat{A}_0$ is zero. In this gauge, the Hamiltonian operator will contain terms

$$\Delta\hat{H} = \int \frac{d^3p_1 d^3p_2}{(2\pi)^6\sqrt{E_1}\sqrt{E_2}} \int d^3x e^{\frac{i}{\hbar}\vec{\mathbf{x}}\cdot(\vec{\mathbf{p}}_2-\vec{\mathbf{p}}_1)} \hat{a}^\dagger_{\vec{\mathbf{p}}_1,s_1} \gamma_0 \vec{\gamma}_{s_1,s_2} \cdot \vec{\mathbf{p}}_2 \hat{a}_{\vec{\mathbf{p}}_2,s_2} + \int \frac{d^3p_1 d^3p_2 d^3p_3}{(2\pi)^9\sqrt{E_1E_2E_3}} \int d^3x e^{\frac{i}{\hbar}\vec{\mathbf{x}}\cdot(\vec{\mathbf{p}}_2-\vec{\mathbf{p}}_1-\vec{\mathbf{p}}_3)} \hat{a}^\dagger_{\vec{\mathbf{p}}_1,s_1} \gamma_0 \vec{\gamma}_{s_1,s_2} \cdot Q\hat{\vec{\mathbf{A}}}^\dagger_{p_3} \hat{a}_{\vec{\mathbf{p}}_2,s_2} + \ldots \quad (381.1)$$

$\hat{\vec{\mathbf{A}}}^\dagger$ is the gauge Boson creation operator, and Q the electric charge. The operator that describes the change taking place in a small instant of time, dt, is $dt\Delta\hat{H}$. The Fermion operators $\hat{a}^\dagger$ and $\hat{a}$ are four-component objects, where the top two components refer to Fermions, and the bottom two to anti-Fermions. Thus we write,

$$\hat{a}^\dagger_{\vec{\mathbf{p}}_1} = \begin{pmatrix} \hat{f}^\dagger_{\vec{\mathbf{p}}_1} \\ \hat{\bar{f}}_{-\vec{\mathbf{p}}_1} \end{pmatrix} \qquad \hat{a}_{\vec{\mathbf{p}}_2} = \begin{pmatrix} \hat{f}_{\vec{\mathbf{p}}_2} \\ \hat{\bar{f}}^\dagger_{-\vec{\mathbf{p}}_2} \end{pmatrix} \quad (381.2)$$

where $\hat{f}$ represents the particle annihilation operator and $\hat{\bar{f}}$ the anti-particle annihilation operator. The γ matrix can connect the creation operator components in each of the spinors.

There is a mathematical relation which states that

$$\int d^3x e^{\frac{i}{\hbar}\vec{\mathbf{x}}\cdot(\vec{\mathbf{p}}_1-\vec{\mathbf{p}}_2)} = (2\pi\hbar)^3\delta^3(\vec{\mathbf{p}}_1 - \vec{\mathbf{p}}_2), \quad (381.3)$$

where δ is the Dirac δ-function which forces $\vec{\mathbf{p}}_1$ (the momentum of the particle) to be equal to $-\vec{\mathbf{p}}_2$ (the momentum of the anti-particle), or $\vec{\mathbf{p}}_1 + \vec{\mathbf{p}}_2 = 0$. In this sum, we add together each momentum corresponding to a creation operator, and subtract any momentum related to an annihilation operator. I will focus on the term with only creation operators, since the principle of causality is equivalent to saying that there should be no such term contributing to the Hamiltonian that describes the change of state. Therefore I add together all the momenta. When we have the gauge field present with momentum $\vec{\mathbf{p}}_3$, we will force $\vec{\mathbf{p}}_1 + \vec{\mathbf{p}}_2 + \vec{\mathbf{p}}_3 = 0$. From Lorentz invariance, the same is true for the energy: the sum of the energies, $E_1 + E_2 + E_3 = 0$ (again in general we add the energies of the particles we are creating, and subtract the energies of those particles we are annihilating, but here we are specifically considering the case of creation from nothing, i.e. no annihilation of particles). But the sum of three positive quantities cannot be zero. Therefore any interaction with only creation operators is forbidden. We need to have an annihilation operator, which means that we need a particle already present for it to act on. Similarly, we cannot have a term made solely from annihilation operators, so we cannot destroy matter and leave ourselves with nothing.

What this means in practice is that the principle of causality holds in the standard model: in physical law, there is no possibility of creating something out of nothing. For that to happen, we would need to go to some agent not bound by physics.

Of course, the standard model excludes gravity. So does gravity change this picture? It is often argued that it does. The point is that gravitational potential energy is always negative (unlike the energy contained in matter which is always positive). So, for example, in Newton's gravity the potential energy between two masses is $-\frac{Gm_1m_2}{r}$; that minus sign means that it is negative. Each lump of matter generates a gravitational field, so when we create a particle and anti-particle, we also create a gravitational potential between them, and therefore some additional negative energy. This, it is argued, would have to be included in the sum over energy, and if we get it just right we might be able to arrange for the energy of the gravitational potential to precisely cancel the mass and kinetic energy of the particles we are creating, allowing the creation of particles out of nothing without violating the conservation of energy.

Of course, before we go to far with this argument, we need to update it to the latest physics. Firstly, in general relativity, gravity is thought of differently from in Newton's theory: there is no gravitational force, and consequently no gravitational potential as Newton conceived it. So arguing from the basis of an analogy from Newton's theory of gravity is a non-starter. Instead, we have the curvature of space time, with particles engaging in inertial motion along geodesics (which may be curved). There is second sense, however, in which we can think of gravity as a negative energy. Einstein's field equation relates the curvature of space time plus the cosmological constant term to the Stress-Energy tensor. Subtract the curvature tensor from both sides of this equation, and we get Stress energy minus curvature equals cosmological constant. In this sense, the curvature of space-time can also be thought of as a negative energy. Create matter, and we curve space time, and maybe that curvature can provide the negative energy that we need to preserve the conservation of energy while creating matter. We need the right circumstances, but maybe we can do it.

However, we should not forget that the definition of energy in quantum field theory is different from that of classical physics, including general relativity. In quantum field theory, it is an index describing the eigenvalue of the Hamiltonian operator for a particular state; in general relativity it is a component of the stress energy tensor. The stress energy tensor can also be constructed in quantum field theory, and it is not the same thing as the energy we are conserving when performing the operator algebra. Only when we take the expectation (or average) values of the fields and take the classical limit do the two coincide.

So let us turn to the calculation used above, and see how that fares when we introduce curved space time. The time evolution operator multiplied by an infinitesimal time slice dt (as it is when we apply the time evolution) is,

$$dt \int d^3x a^\dagger_{s_1}(\vec{\mathbf{x}})\gamma_0\vec{\gamma}_{s_1,s_2} \cdot (i\hbar\vec{\partial} - Q\vec{\mathbf{A}}(\vec{\mathbf{x}}))a_{s_2}(\vec{\mathbf{x}}). \tag{382.1}$$

When we apply this in curved space time, we need to replace the derivative operator with the covariant derivative, $\vec{\nabla}$, and adjust the measure of the integral by the square root of the determinant of the metric tensor $g_{\mu\nu}$, giving,

$$dt \int d^3x \sqrt{-\det(g_{\mu\nu})} a^\dagger_{s_1}(\vec{\mathbf{x}})\gamma_0\vec{\gamma}_{s_1,s_2} \cdot (i\hbar\vec{\nabla} - Q\vec{\mathbf{A}}(\vec{\mathbf{x}}))a_{s_2}(\vec{\mathbf{x}}). \tag{382.2}$$

We can try to apply the same procedure as before: replace $\hat{a}$ with $\hat{a}e^{\frac{i}{\hbar}\vec{\mathbf{p}}\cdot\vec{\mathbf{x}}}$ and use equation (381.3) to produce a Dirac-δ function that enforces the conservation of momenta. But now we encounter a problem: both g and $\vec{\nabla}$ are functions of the location x. Thus we do not obtain the same integral over location as in (381.3), which means that we do not get a Dirac-δ function forcing the conservation of momentum and energy, so it seems as though the argument above breaks down.

The flaw in this reasoning was that it assumed that the eigenstates of the Hamiltonian are the same in flat and curved space time (i.e. corresponding to $e^{\frac{i}{\hbar}\vec{\mathbf{p}}\cdot\vec{\mathbf{x}}}$). Instead, we should use the eigenvectors of the covariant derivative $\vec{\nabla}$. Recall that $\vec{\nabla}$ can be written as the derivative operator in an alternative coordinate system, $\nabla_i = \tilde{\partial}_i = \frac{\partial}{\partial\tilde{x}_i}$. The eigenstates[12] of the derivative operator are thus $e^{\frac{i}{\hbar}\tilde{\vec{\mathbf{p}}}\cdot\tilde{\vec{\mathbf{x}}}}$. We can rephrase the integral in terms of the new $\tilde{\vec{\mathbf{x}}}$ variables, which gives

$$d\tilde{t} \int d^3\tilde{x} a^\dagger_{s_1}(\tilde{\vec{\mathbf{x}}})\gamma_0\vec{\gamma}_{s_1,s_2} \cdot (i\hbar\tilde{\vec{\partial}} - Q\vec{\mathbf{A}}(\tilde{\vec{\mathbf{x}}}))a_{s_2}(\tilde{\vec{\mathbf{x}}}), \tag{382.3}$$

[12]There are some complications here, because the derivative operators in curved space time don't commute with each other when acting on a vector, which means that no vector can be a simultaneous eigenstate of derivative operators in curved space time. However, the operators are torsion free, which means that they do commute when acting on a scalar such as $e^{\frac{i}{\hbar}\tilde{\vec{\mathbf{p}}}\cdot\tilde{\vec{\mathbf{x}}}}$. Thus we can express this as a simultaneous eigenfunction of each of the momentum operators, even in curved space time.

and now by expanding the creation operators in terms of the quantities parametrising the eigenstates in the curved space time, $a_{s_2}(\tilde{\vec{\mathbf{x}}}) = \int \frac{d^3\tilde{p}}{(2\pi)^3\sqrt{\tilde{E}}} a_{s_2}(\tilde{\vec{\mathbf{p}}}) e^{\frac{i}{\hbar}\tilde{\vec{\mathbf{p}}}\cdot\tilde{\vec{\mathbf{x}}}}$. Following this substitution, the argument proceeds as before.

So rather than generating a gravitational field with a negative energy, the effect of including gravity is to alter the energy eigenstates. The eigenstates, however still have to be positive, for the same reasons as in flat space-time. Thus, even in curved space time, it is still impossible to create something out of nothing through physical processes.

What if there is other matter in the universe? Won't the deformation of space time lower the energy levels of that matter, allowing for the spontaneous creation of matter and the conservation of energy? But then the presence of that other matter violates the assumption that there should be nothing in the initial state interacting with our created particles; the other matter (or black hole, or whatever) would be the efficient cause of the creation, mediated through gravity.

Another issue with the spontaneous creation of matter is that it will lead to a discontinuous change in the space time metric. The ordinary decay of matter doesn't do so, because the energy and momentum are conserved at the moment of decay (at least up to minor quantum fluctuations, which average out over the large scale that affects gravity – recall that in the context of gravity we need to discuss the stress energy tensor rather than the quantum definition of energy and momentum). The changing pattern of the way in which that energy is distributed will lead to a change in the metric, but a continuous and gradual one. Continuous change is fully consistent with the principles of general relativity; discontinuous change in the metric is more troubling, and will lead to infinities in the Christoffel symbols and curvature tensor, and consequently mathematical inconsistencies.

Of course, this analysis has the weakness that it was based on quantum field theory in curved space time, rather than a true theory of quantum gravity. But with spontaneous creation impossible both in quantum field theory, and classical general relativity, due to the conservation of energy, why should we expect it to be different in theory that combines the two? ■

Discussion:
This is the key premise we shall need in subsequent sections while discussing the existence and nature of God. I have claimed that, while undermined by classical physics, this is fully consistent with and required by quantum physics, given that in quantum physics the state of an object can be defined in terms of the eigenstates of the Hamiltonian, that every free particle is in a singe eigenstate (rather than superposition) of the Hamiltonian, and that every particle in an eigenstate of the Hamiltonian remains in that state until it either decays or absorbs another particle. The proof that you cannot get something from nothing follows from the conservation of energy and momentum.

I should emphasise again that while I have confirmed the principle from an argument from physics, the principle for causality is more fundamental than physics: it is a requirement for reality to be rational and understandable. □

Objection:
A radioactive nucleus (for example) decays spontaneously; there is no possible way to predict when it will decay, and no external cause that makes it decay. Thus this decay is either caused by the particle itself or is uncaused. Given that something cannot cause itself, this interaction must be uncaused. Therefore the principle of causality is violated. •

Response:
This objection has been raised against arguments in favour of the principle of causality

which rely on the idea that if an event did not have a cause, we would not be able to predict when it would happen. However, although this argument is invalid, the principle of causality does not depend on it, but may be proven by considering the various conservation laws in nature.

At the risk of being slightly too simplistic, let us suppose that a particle described by a state $|A_1\rangle$ has only two possible interactions within a given time period: it may decay into another state (say $|B_2\rangle$, which describes two particles), or continue in its current state. Thus there are two possible outcomes within a given time period:

$$\begin{aligned} |A_1\rangle &\rightarrow |A_1\rangle \\ |A_1\rangle &\rightarrow |B_2\rangle \,. \end{aligned} \tag{384.1}$$

In the first case, there is no change, so the principle of causality does not apply. In the second place, there is change, but we can say that $|A\rangle$ is the efficient cause of the particles contained within $|B_2\rangle$, while $|B_2\rangle$ is a final cause of $|A\rangle$. Initially, $|B_2\rangle$ is a potentiality, and it is actualised by $|A\rangle$, which is initially actual (although it is not at the end). Therefore the principle of causality is not violated.

To go into more detail, we need a fuller discussion of the physics of spontaneous decay. It is possible to compute the decay rate of any system; the calculation is the same in both quantum mechanics and quantum field theory. The final result is known as Fermi's golden rule, and I offer a brief derivation of it below.

We wish to consider the process where an initial state $|\psi_i\rangle$ decays into a final state $|\psi_f\rangle$, where $|\psi_i\rangle \neq |\psi_f\rangle$. Decay is possible when these are not eigenstates of the Hamiltonian. This presents us with an immediate conceptual problem – the eigenstates of the Hamiltonian represent the possible potentia of the system; so can $|\psi_i\rangle$ and $|\psi_f\rangle$ be defined? The key is that we consider a Hamiltonian operator that changes in time. At the initial time and final time, the system is governed by a Hamiltonian $\hat{H}_0$; and we use this to define the initial and final states. At intermediate times, we add to this a small perturbation $\hat{H}_1$. Perhaps the system changes because a second particle approaches the one which is going to decay (so the form is defined in terms of the Hamiltonian that contains an average background for all the other matter in the system; and $\hat{H}_1$ refers to the deviations from this average), or an experimenter flicks a switch. So between the initial and final states, we consider a Hamiltonian $\hat{H}_0 + \hat{H}_1$; and the initial state is not in an eigenstate of this Hamiltonian, so decay is possible. For this derivation I will treat $\hat{H}_1$ as though it is constant in time; the argument can be extended to cover a changing $\hat{H}_1$, but this just adds additional complications without affecting the final conclusion I want to draw.

As a concrete example, let us consider the case when an electron emits a photon (a quanta of electromagnetic radiation). $\hat{H}_0$ may represent the (renormalized) free particle Hamiltonian, which describes matter without any interactions, so electrons, photons and so on continuing without any change in momentum; or it might describe the orbit of an electron around a positively charged nucleus (for example, this might indicate a change of an exited electron to the ground state). $\hat{H}_0$ describes a system without interactions; and this is the operator we use to identify meta-stable states, including our initial and final states. $\hat{H}_1$ represents an interaction term, perhaps between a free electron and an external electromagnetic field, or between an atom and a laser beam an experimenter has just switched on. $\hat{H}_1$ can cause a change in the state; in this case either the absorption or emission of a photon.

At times very far in the past and very far in the future, the system is in an eigenstate of the Hamiltonian $\hat{H}_0$. We label these eigenstates as $|\psi_n\rangle$ with energy eigenvalues E_n. Obviously, it remains in this eigenstate over time. Then the Hamiltonian changes and we start the clock. The particle is no longer in an eigenstate of the Hamiltonian which is now $H_0 + H_1$, with new energy eigenvalues E'_n with eigenvectors $|\phi_m\rangle$. We assume that the perturbation $\hat{H}_1$ is small (one can, of course, represent a large perturbation as a large

number of small perturbations, so this assumption need not restrict the applicability of my argument, but just makes it simpler). Instead the state of the particle will be in a linear combination of the eigenstates of this new Hamiltonian, and the factors of e^{iEt} in front of each of the new energy states will evolve differently in time. If we compare this new state with the eigenstates of the original Hamiltonian, the coefficients of the expansion (recall that they represent the likelihood that when measured the particle will be found in a particular state) will gradually change with time. We can write that

$$|\psi_n\rangle = \sum_m \alpha_{nm} |\phi_m\rangle , \tag{385.1}$$

for some coefficients α. I order the eigenstates so that $|\psi_n\rangle$ evolves into $|\phi_n\rangle$ as we gradually switch on the perturbation $\hat{H}_1$. Now $|\psi_n\rangle$ contains information concerning both the electron and the photon, and to keep this explicit I will write it as $|\psi_n\rangle = |e_n\rangle \otimes |\gamma_\omega\rangle$ to describe an electron in state n with energy E_n and a photon γ with energy $\hbar\omega$; I will use $|0\rangle$ to describe the state where there are no particles of that type present, and for simplicity (since this is all I need) I will assume that at most there is one particle of each type.

The initial state contains one electron and no photons, which can be described by $|\psi_i\rangle = |e_i\rangle \times |0\rangle$, which is an eigenstate of $\hat{H}_0$ with energy E_i. The final state contains an electron and a photon, $|\psi_f\rangle = |e_f\rangle \times |\gamma_\omega\rangle$, an eigenstate of $\hat{H}_0$ with energy $E_f + \hbar\omega$ (E_f represents the energy of the electron, and $\hbar\omega$ the energy of the photon). The evolution in time is given by the eigenvalues E' of $\hat{H} = \hat{H}_0 + \hat{H}_1$, so at time t the system will be in the state

$$|\psi(t)\rangle = \sum_m \alpha_{im} e^{iE'_m t/\hbar} |\phi_m\rangle . \tag{385.2}$$

The likelihood that when we measure the energy of the particle using the Hamiltonian $\hat{H}_0$ that we will find it in the state $|\psi_f\rangle$ is thus

$$\langle\psi_f|\psi(t)\rangle = \sum_m \alpha_{im} e^{iE'_m t/\hbar} \langle\psi_f|\phi_m\rangle . \tag{385.3}$$

At this point, I make the assumption that the energy states all have different eigenvalues (again, we need not assume this, and relaxing the assumption makes no difference to my overall conclusion but only complicates the argumentation, so for this presentation I will keep things simple by making this assumption), and that the interaction $\hat{H}_1$ is sufficiently small that we can neglect terms of $O(\hat{H}_1^2)$ (again, we can construct the argument without this assumption, but it just adds unnecessary complications without affecting my conclusions). $\hat{H}_1$ is constant in time; the interaction between electrons and photons does not change in time. Now if $\hat{H}_1$ is small, then the Hamiltonian $\hat{H}_0 + \hat{H}_1$ will be similar to the Hamiltonian $\hat{H}_0$ and consequently the eigenstates $|\phi_n\rangle$ will be similar to the eigenstates $|\psi_n\rangle$. In fact, we can compute what the difference between them is, and it turns out that

$$|\phi_n\rangle = |\psi_n\rangle + (1 - |\psi_n\rangle\langle\psi_n|) \frac{1}{\hat{H}_0 - E_n} \hat{H}_1 |\psi_n\rangle + O(\hat{H}_1^2), \tag{385.4}$$

where the $O(\hat{H}_1^2)$ indicates that there are terms which contain the square and higher powers of $\hat{H}_1$ which we have omitted. This leads to

$$\alpha_{im} = \langle\phi_m|\psi_i\rangle = \frac{\langle\psi_m| \hat{H}_1 |\psi_i\rangle}{E_i - E_m}(1 - \delta_{im}) + \delta_{im} + O(\hat{H}_1^2) \tag{385.5}$$

$$\langle\psi_f|\phi_m\rangle = \frac{\langle\psi_f| \hat{H}_1 |\psi_m\rangle}{(E_f + \hbar\omega) - E_m}(1 - \delta_{fm}) + \delta_{fm} + O(\hat{H}_1^2) \tag{385.6}$$

and, given that $\langle\psi_f|\psi_i\rangle = \langle e_f|e_i\rangle\langle\gamma_\omega|0\rangle = 0$ since there is no overlap between the initial and final photon states, we have

$$\langle\psi_f|\psi(t)\rangle = \frac{\langle\psi_f|\,\hat{H}_1\,|\psi_i\rangle}{E_f+\hbar\omega-E_i}(e^{iE_it/\hbar} - e^{i(E_f+\hbar\omega)t/\hbar}) + O(\hat{H}_1^2), \tag{386.1}$$

which comes from the sum of the contributions when m is taken to be the initial state i and where m is taken to be the final state f (which has energy $E_f+\hbar\omega$) – the remaining possibilities for m give contributions of $O(\hat{H}_1^2)$, so we can neglect them. The probability of finding ourselves in state f at time t is given, as always, by the square of the likelihood,

$$P(f\setminus i;t) \equiv |\langle\psi_f|\psi_i\rangle|^2 = \frac{4\sin^2((E_f+\hbar\omega-E_i)t/2\hbar)}{(E_f+\hbar\omega-E_i)^2}|\,\langle\psi_f|\,\hat{H}_1\,|\psi_i\rangle\,|^2. \tag{386.2}$$

This gives the probability of a transition to a final state $\psi_f = |e_f\rangle\otimes|\gamma_\omega\rangle$ with electron energy E_f and photon energy $\hbar\omega$. To calculate the total probability of an electron emission, for all possible final states, we need to multiply by a density of states $\rho(E)$ and integrate over the possible final energy states.

$$P(\text{decay}\setminus i;t) = \int dE_f\rho(E_f)\frac{4\sin^2((E_f+\hbar\omega-E_i)t/2\hbar)}{(E_f+\hbar\omega-E_i)^2}|\,\langle\psi_f|\,\hat{H}_1\,|\psi_i\rangle\,|^2, \tag{386.3}$$

and at times where $d/dE_f(\sin^2((E_f+\hbar\omega-E_i)t/2\hbar)) \gg d/dE_f\rho(E_f)$, we solve this equation[13], and we finish with Fermi's golden rule

$$\frac{dP}{dt} = \sum_f \frac{2\pi}{\hbar}|\,\langle\psi_f|\,\hat{H}_1\,|\psi_i\rangle\,|^2\rho(E_i-\hbar\omega)\delta(E_f+\hbar\omega-E_i). \tag{386.4}$$

where we sum over all final states consistent with the energy conservation implied in the δ-function. The rate of decay therefore depends on the matrix element $|\,\langle\psi_f|\,\hat{H}_1\,|\psi_i\rangle\,|^2$. The allowed decays are those options where this element is non-zero; this will rule out possible decay paths where spin, electric charge and so on are not conserved.

We now need to review what we have just computed. $|\psi_i\rangle$ represents the initial, actual state, $|\psi_n\rangle$ the potentia, and $|\psi_f\rangle$ the potentia that is actualised in the final state. The possible final causes of $|\psi_i\rangle$ are encoded in the interaction Hamiltonian $\hat{H}_1$. Now $\hat{H}_1$ tells us the rules governing the emission or absorption of a photon by an electron; acting on a vacuum state it gives nothing. The Feynman rules tell us that the matrix element $\langle\psi_f|\,\hat{H}_1(|0\rangle\otimes|0\rangle) = 0$, which in words means that if the initial state contains no particles, it cannot be changed by $\hat{H}_1$. This is quite general, and applies to every physical system, since the Hamiltonian operator always has the structure creation operator times something times annihilation operator[14] and the annihilation operator on a state with no particles always gives zero. If we were to include higher orders of $\hat{H}_1$ in the calculation, then the corrections to Fermi's golden rule resemble $\langle\psi_f|\,\hat{H}_1\,|\phi_{m_1}\rangle\,\langle\phi_{m_1}|\ldots|\phi_{m_n}\rangle\,\langle\phi_{m_n}|\,\hat{H}_1\,|\psi_i\rangle$, and we still have $\hat{H}_1$ acting on ψ_i, so if the initial state had no particles we would still get zero. In other words, one cannot have a transition to a final state containing a non-zero number of particles from an initial state with no particles in it. A potentia cannot be actualised unless the initial state is actual. Secondly, not every possible final state is permitted, but only those which overlap with the initial state according to the Feynman rules contained within $\hat{H}_1$ (we have already seen the conservation of energy above, there are also rules governing spin/angular momentum and charge conservation as well); this is equivalent to the principle of proportionate causality.

The decay of a particle is a problem for certain interpretations of traditional Aristotelianism, where it is implied that for $|e_i\rangle$ to change into $|e_f\rangle$ would require the intervention of

[13]Integrate by parts then perform a contour integral picking up the residue of the pole at $E_f+\hbar\omega-E_i$.

[14]This is complicated when we have anti-particles, since the particle creation operator is the anti-particle annihilation operator. However, in all the relevant terms in the Hamiltonian there would also be a photon or Higgs annihilation operator, which would give zero when applied to the vacuum state.

something external to the electron, which is not the case here. But in this formulation, perhaps a slight modification of the traditional one, there is no issue with the principle of causality. ○

Objection:
In quantum field theory, people often talk about pulling particle/anti-particle pairs out of the vacuum. In this case, something arises from nothing. This violates the principle of causality above. •

Response:
The Feynman diagram approach allows for the possibility of closed electron or gluon lines; these represent particles which are pulled out of the vacuum. However, they do not interact with anything; their contribution is explicitly removed in the quantum field theory analysis; from the definition of existence (4.5.1) they do not exist, as nothing depends on their presence.

Alternatively, some discuss pulling particles out of the vacuum in the sense of an quark/anti-quark pair arising from a virtual photon (for example). This, however, does not violate the principle of causality, as the photon is already actual. In this case, we are better saying that the photon decays into a particle/anti-particle pair.

The conservation of four momentum states that a particle and anti-particle cannot annihilate each other and produce nothing – since the particles have mass, they cannot have in total zero energy, and that energy has to go somewhere. By time reversal symmetry, the opposite must also be true: the particles cannot be created from nothing without a violation of the conservation of energy and momentum. Heisenberg's uncertainty principle (in the form $\Delta E \Delta t \geq \hbar/2$) is often invoked as a way around this: there can be a small violation of energy conservation if the time uncertainty Δt is small enough. But this is a misuse of the principle: this form of the uncertainty relation describes the width of the peaks of the time and energy wavefunctions of a particle that is already actual. It is not at all clear that it applies for the wavefunctions of fields in a vacuum state where there is nothing actual.

Therefore, this objection does not invalidate the principles of causality as long as energy and momentum are conserved, i.e. translation invariance of space-time and the locality of the action are maintained. ○

CHAPTER 12

Modern Objections to A Classical theory of God

12.1. Introduction

In the following chapters, I shall review and argue for the understanding of God prevalent in classical theism. It should be emphasised at the outset that this differs from many contemporary visions of God, particularly the rather simplified view held by most 'men on the street.' This includes most atheists and a fair proportion of the churchgoers, even those who spend their time in the more orthodox churches which, in principle at least, still hold to a classical conception of God. However, most of the major Christian scholars of the past were some variety of classical theists. Of course, there is variation among classical theists; I make no claim that every major Christian thinker would have agreed with every detail of my presentation. In particular, we can list three main divisions. The earliest Christian theologians, and the Eastern Orthodox to this day, adopted neo-Platonism. In Latin Christianity, there was a drift towards a larger emphasis on Aristotle's philosophy. The more philosophically minded Protestant reformers (in whose tradition I stand) built on the Latin view, but trimmed away what they regarded as the wilder speculations to return to what they saw as a more authentic representation of the original Christian belief. But, at least as far as the doctrine of God is concerned, there is very little practical difference between these three groups; and I think most people in them would agree with my conclusions, if not necessarily all of my premises and how I get there. It is only in the past few centuries that other approaches to theism have become more prominent. These include the more radical evangelicals who largely ignored philosophical theology; the modernists who adopted enlightenment philosophy, and are generally scattered from deism to vague spiritualism to views which are close to atheism; and, more recently, notable Christian Philosophers such as William Lane Craig and Alvin Plantinga who are theists, but disagree with the classical conception of God. Most atheist philosophical works I have read (there are a few notable exceptions) do not engage with the classical theist tradition. To pick one example, I can read Makie's *The Miracle of Theism* (?), one of the most important atheist texts, and agree with most of what he wrote, aside from his chapter recapitulating Hume's argument against miracles, his rather superficial discussion of the cosmological argument, and, of course, his overall conclusion. He simply doesn't address the classical picture of God, but mainly responds to views that I myself don't accept either.

If you are not a knowledgeable classical theist, it is therefore important that you put any pictures of God you already have out of your mind while reading this, as they will only serve to mislead you. The description is, however, fully consistent with Christian tradition. But since classical theism rather than Christianity is the focus of this work, I shall not draw out any specifically Christian doctrine; other monotheists, such as Jews and Muslims, will also find much in these pages which they can agree with. I must emphasise that this is not the only version of philosophical theism consistent with Christian revelation; to my mind, however, it is the one that explains the most.

These chapters are not intended to be a full description of classical theism (which would require several volumes rather than a few chapters); however, I intend to argue for and describe the main points using the premises and definitions already established, and to show the whole picture's self-consistency. The strongest evidence for Christianity, I believe, lies

away from these arguments and from observed data, but without the background provided here it is difficult to appreciate and accurately evaluate that evidence.

I intend to discuss the definition, existence and properties of God using, as my starting point, the premises discussed in the earlier chapters, particularly causality. My presentation, arguments and approach are inspired by those of Thomas Aquinas, although I will differ from his work in certain aspects. My purpose is to present a framework which can be used to evaluate and understand the evidence for God. Many people come to the evidence with a preconceived understanding of nature. Hume famously argued against the possibility that we should accept an account of miracles, but his argument was based on his metaphysical presumptions, which are unconsciously shared with and accepted by many modern atheists without sufficient reflection or examination. I believe that Hume's metaphysics is incompatible with theism; I also firmly believe it to be wholly and completely wrong, as reinforced by its various tensions with the experimentally derived quantum physics.

Most atheists try to fit God into their own view of metaphysics without any other changes, which doesn't work. But they are not approaching the evidence in an unbiased manner. They approach the evidence assuming that certain metaphysical propositions are true. What if those propositions are false? What if there is some other way of understanding nature which is consistent with, or even demands, the miraculous? To understand God, one must also also understand a world-view which is very different from our modern way of thinking. Before we can come to Christianity, or any other revealed religion, we have to understand what theists mean when they speak of God. My purpose in these next chapters is to explain an understanding of the world and of God which is consistent with both theism and with modern physics. In short, to introduce a way of thinking which allows one to fairly judge the evidence, so that we have it available to us when looking at the observed data. We do not assume that the presuppositions of the modern world are correct; but leave ourselves open to other possibilities.

If you have found a paradox in a statement made by a classical theist, I claim that it is not in theism itself that the contradiction lies, but because you have tried to interpret that statement through some other paradigm, such as mechanism, which is incompatible with classical theism. First one needs to understand theism on its own terms; then one can properly judge and evaluate it.

12.1.1. Knowability of God

The first objection is what right have I to define God? Am I not trying to constrain the unconstrainable, or know the unknowable?

Many modern scholars believe that we cannot know or understand God[1]. At best we get a partial glimpse. In short, although there is a divine reality out there somewhere, it is infinite and our finite minds simply can not grasp its entirety. Some people get a partial glimpse of a divine presence, as the barriers between us and God are weakened, and what

[1]Schleiermacher, one of the main influences of liberal theology, is one notable example of this trend. In his Christian faith, Schleiermacher defined God in terms of a *feeling of absolute dependence*, and wrote *The recognition that this feeling of dependence is an essential condition of life takes for us the place of all proofs of the existence of God ; which proofs have no place in our procedure* (Schleiermacher, 1884). Later on, he wrote, *All attributes which we ascribe to God are not to be taken as indicating something specific in God, but only something specific in our manner of referring to Him the feeling of absolute dependence* (Schleiermacher, 1884). This implies that nothing positive can be known about God; everything we say about God relates to our own feelings and thus is only subjective. It will not surprise the reader that I reject almost everything that Schleiermacher wrote: God is far more than just a feeling, and reducing our relationship with God to mere feelings makes it worthless and incapable of providing firm evidence. God exists objectively and the evidence for His existence must be open to objective examination or He is nothing at all. Schleiermacher's overall philosophy is derived from Kant, and contains Kant's fundamental flaws, plus a few more of his own invention. My own thought is that his version of Christianity, and those inspired by him, are not intellectually tenable from a scientific perspective; however this is not the place for a detailed criticism of his writings.

they perceive is is useful for them but not necessarily useful for those in a different context. An analogy which is frequently cited is that of the blind men examining an elephant: each feel a different part of it, and come to a different conclusion as to what it is; and they subsequently have a sharp disagreement concerning its nature. None of them are wrong, and none of them are fully right. The truth is simply bigger than any of them are able to understand.

This understanding is incompatible with Christianity. There are two aspects of this. Firstly, is God knowable through revelation? Secondly, is God knowable through reason? To the first, every Christian has to answer 'yes':

> And the Word became flesh and dwelt among us, and we have seen his glory, glory as of the only Son from the Father, full of grace and truth. ... For the law was given through Moses; grace and truth came through Jesus Christ. No one has ever seen God; the only God, who is at the Father's side, he has made him known[2].

This passage, which is just one example I could have quoted, states that through Jesus, God has been made known. If God has been made known, then he is knowable: we can know facts about him. To deny that God can be known is either to deny that we can know Jesus, or to deny the incarnation. Secondly, we may claim that God is knowable through revelation, but not by reason. Again, this is against Christian testimony,

> For the wrath of God is revealed from heaven against all ungodliness and unrighteousness of men, who by their unrighteousness suppress the truth. For what can be known about God is plain to them, because God has shown it to them. For his invisible attributes, namely, his eternal power and divine nature, have been clearly perceived, ever since the creation of the world, in the things that have been made. So they are without excuse. For although they knew God, they did not honour him as God or give thanks to him, but they became futile in their thinking, and their foolish hearts were darkened. Claiming to be wise, they became fools[3].

Again, it is clear that that this is saying that the attributes of God can be known by observing the world[4]. Not that they can be partially known, or we can discover something that can be true for us but not for others. Paul's argument is that everyone falls short of God's standard, and that ignorance is not an excuse, since we have enough evidence available to us to discern enough about that standard to stand condemned. Equally, since the evidence is universal, it suggests that the standard is universal (a point abundantly affirmed elsewhere in the Bible); it is not that one person examines the tusks of the elephant and another the tail and that their differing conclusions are equally valid: they each have the opportunity to examine both the tusks and the tail, and after that if they disagree then at least one of them is wrong. If the standard is subjective then Paul's conclusion that all stand objectively condemned[5] does not follow.

This does not mean that the modern position is wrong; merely that if you believe it then you must believe that Jesus and those He appointed to convey His message were mistaken:

---

[2]John 1:14-18.

[3]Romans 1:18-21.

[4]Note that I am saying that it *can* be known; it does not mean that it *will* be known, or that God's existence is self-evident to everyone. It means that if one searches well enough according to the right presumptions, then one would have to conclude that God exists. Atheists are obviously not all fools (or no more foolish than anyone else), but, from this perspective, either proceed from false presumptions, or by not putting enough effort into the search, or are distracted by bad arguments for God whether straw-men raised by those such as Hume or nonsense as spouted by the likes of Kant and Schleiermacher, or swayed by apparent evidence against God (valid arguments against various misconceptions of God, but not the true God). It is not deliberate rebellion, but a mistake in evaluating the evidence, or spending too much time dwelling on the wrong evidence.

[5]Romans 3:23

in short, you cannot be a Christian. For the Christian asserts that we can know God, both through God's self-revelation (God reaching out to us and telling us about Himself and His nature) and reason (our investigation of nature).

The modern position is unpalatable from the perspective of a scientist as well. To claim that nothing can be known for certain, that everything is subjective about something which must, by its very nature, be objective and universal runs against the assumptions on which contemporary science is based. To say that we do not know is perfectly acceptable; to say that we cannot know is troublesome, to claim that God is ultimately unknowable is to explain why many modern theologians and philosophers of religion are held in contempt by most scientists (including myself). If God is the origin of scientific law, and scientific law is objective, then God cannot be subjective, since a subjective premise cannot give rise to an objective conclusion. If God is not knowable, then no conclusions can be drawn from trying to study Him, and religious practice is built solely on custom and assertion, and anything which is not founded on objective truth is just a sham. If nothing can be known about God, then we know nothing about God, and He may as well not exist for all the good our ceremonies based on ignorance would do.

The question then becomes 'To what extent is God knowable?' Now, by asserting that God's nature is knowable by reason I am of course assuming that God is rational; that is that He is self-consistent and cannot be of another form than His own, or simultaneously two different things at once. This assumption is necessary for any inquiry. It is implied by the observation that the universe is rational; if God is the source of the universe, then it is strongly unlikely that an irrational God would give rise to a rational universe.

Now, I am not claiming that we can know God down to every last detail, any more than we can understand the matter contained within the universe down to every last detail (for example, to name every speck of space dust in a solar system in a distant galaxy). We can know of God with certainty within certain bounds (that is to say that God must be this, this or this), and in particular we can know that certain statements about God are false. Nor am I constraining God by making these deductions: I am describing Him. I do not constrain an electron by asserting that its rest mass must be within a certain range; I describe the properties that an electron has been observed to possess. If God is discernible to the eye of reason, then we should not be surprised that our reason will tell us that God possesses certain attributes and not others. My description may be wrong; but since everything that exists must exist as something, or must have a particular essence or nature; that nature can be described and understood, if not fully then at least to a certain extent: there must be some correct description of God (if He indeed exists). This where the argument that our finite minds cannot comprehend the infinite God fails: our finite minds may not be able to objectively comprehend all of infinity, but that is no reason to suppose that we can objectively comprehend none of it.

It is foolish then, to start our inquiry with the presumption that we cannot succeed. We may fail due to our limitations; but that is something to be seen after the investigation but not assumed before it, and if we do fail, then the best response is simply to try again. Let us investigate, and see how far we can go without making any prior assumption about how far that will be.

12.1.2. Use of Metaphysics

The second objection is am I not, by using an Aristotelian-Thomistic metaphysics, trying to fit reality to that metaphysics, rather than fit the metaphysics to reality? Am I repeating the error of the Traditional Chinese Medics, who tried to shoehorn their worthy empirical investigations into a philosophy of Taoism and dualism, even though, despite a few successes, the fact is that Taoist vitalism simply does not provide an accurate description of how the human body works? They try to fit their observations into metaphysics, rather than to judge their metaphysics by observation and discard it when it proves inadequate?

Certainly, were this true, it would be a serious objection. We are, however, perfectly justified in adopting a model if it is true; that is that it is in agreement with objective fact. Before we can discern whether something is correct, we must understand it, justify its premises, and construct predictions which can then be tested. But it is impossible to even start a rational inquiry without adopting some metaphysical model to provide the foundational axioms and premises. The wrong is not in adopting a model; the wrong is holding to that model when it has been tested and proved false. If it fails the test, then it must be either modified or abandoned in favour of some other model – just as the Chinese theory that disease is caused by the imbalance of yin and yang in the body (for example between heat and cold, or moisture and dryness), or a disruption in the flow of blood or the vital energy qi must be surrendered in favour of the description of modern medicine in terms of viruses, fungi, bacteria, cancers, and damage to and degradation of the body's cells.

However, that the Chinese model of the body is wrong does not mean that every model of the body is wrong. There must be one right answer, and modern medicine and biochemistry is well on the way to finding that correct model. Similarly, with respect to metaphysics, we should not conclude from the failure or uselessness of some metaphysical models (such as the mechanical world view that arose in the enlightenment, or most contemporary approaches to metaphysics) that all metaphysical models are wrong. Instead, we should press on with the model in closest conformity to modern physics, experiment, experience and simple common sense, which to my mind is by far the scholastic model. Should this metaphysical model be found to be incorrect – and undoubtedly it will need to be fine tuned and perfected at the very least – then it must be modified or abandoned. But before we decide whether we ought to adopt, modify or abandon it we must first examine it and its consequences.

I certainly do not wish to claim that the metaphysics I am adopting here is the only way of understanding classical theism. There is, for example, a strong stream of Christian theology based on the ideas and work of Plato (which continues to be practised by many Eastern Orthodox theologians today), and there are other approaches neither Aristotelian nor Platonic[6]. Even within the general Aristotelian framework, there are differences and various options. I have chosen the system I have firstly because I am more familiar with it than the others; secondly because both physics and theism seem to me to fit into this framework easily; thirdly because it is well studied by others, and there is a large body of work for me to draw on, and its theological implications are well understood and tested and found to be consistent with orthodox Christian theology. I do not claim that other approaches to physics and theism are wrong: they may fit the data well, or perhaps even better than what I present here. The methods may also be equivalent, in that they reach the same overall scheme but by taking some of my conclusions as their axioms, and deriving my axioms as their conclusions. Nor do I claim that I have every detail correct; there are doubtless areas where my arguments can be improved, and it should certainly not be thought that because I have made a mistake here that theism itself or my entire project falls down. If I have failed, then there are other approaches which may well succeed. Nonetheless, given my understanding and knowledge, I believe this to be the most fruitful of the available approaches.

12.1.3. Ontological arguments

The third objection is that am I by adopting philosophical arguments for God, simply playing with words? A conclusion is no better than its premises, and by trying to argue on the basis of metaphysics the best I can demonstrate is that my view is self-consistent, not

[6]Indeed the earliest Christians were not philosophical at all, and some, such as the Baptists and free churches, follow in this tradition. I believe that approach is flawed, because only if a philosophy is inconsistent with empirical observation and Christian revelation should a Christian reject it; some philosophy is correct, and to reject all philosophy restricts our thinking. It does, however, protect against the opposite error of placing too much of an emphasis on philosophy, and letting bad philosophy introduce error into our religious understanding.

that it is a good description of the world as it is. This objection is often raised against the ontological argument for God (which I shall not use), and it is sometimes claimed that the cosmological argument falls into the same trap.

Those who argue this hold to a view that knowledge must at least in part be based on observation; that nothing can be deduced through pure reason alone. I am not convinced by this; there are areas within logic and mathematics where we can say that given the definitions the conclusion must be correct, and it is foolhardy to say that all that emerges from mathematics are the definitions which we put in[7]. A mathematical argument can be objectively true, even if it is not observable empirically or even demonstrated rigorously apart from some external axiom.

But this is not what I am advocating here. I will make use of various forms of the cosmological argument for God. It should not be claimed, as was by no less a figure than Immanuel Kant, that the cosmological argument is purely deductive. It is built on three premises: substance causality, the impossibility that the sequence of causes that led to us can extend infinitely into the past, and that something exists. There are purely philosophical arguments for the first two premises, but all of them can be, and the third has to be, justified and tested empirically. They are all consistent with our present knowledge of physics, and should only be rejected as implausible if our understanding of physics changes.

Equally, the arguments here lead to a world view which makes certain predictions (or postdictions), most particularly about the possibility of the miraculous. These can be tested by observation; not scientifically (as it is impossible to run a controlled experiment which has an uncontrollable God as one of its variables), but not every observation is scientific. In short, my argument is not ontological: it is derived from, and can be tested by observation.

12.1.4. Does this prove what I need?

The fourth objection is that even if I succeed, is my argument of any use to Christianity? Could it not equally be applied to some other non-Christian religion such as Hinduism, Islam, Episcopalianism, Buddhism, the cult of the Airborne Pasta Beast, Sith Lords, Unitarianism or any other religion conceived, unconceived or inconceivable?

This objection would have some force if I was trying to prove Christianity through reasoned argument. I am not. I am not claiming that I can derive all of Christianity through this argument; merely that I am constructing a framework in which Christianity makes sense, along with some (though not all) of the other religions I mentioned. At the centre of Christianity is the crucifixion and resurrection of Christ; it may, perhaps, be deduced that something similar to that is necessary, but not that it occurred in that form or in that time. To describe Christianity we require both reason and observation: reason alone (meaning reason based on the findings of physics apart from any special revelation) is not enough.

However, reason alone is powerful enough to describe much of classical theism. We may not reach any particular revealed religion, but we will be able to say that if a religious belief (including atheism) describes God or the gods in a certain way then that religion is false, if our premises hold. We will not prove Christianity (for example), but we will eliminate many of the alternatives. My purpose here is not to prove a religion through reason, but to construct a framework within which revelation can be understood and against which it can be tested; so we can approach the evidence in a manner which allows us to be fair to it.

---

[7]Of course, Gödel's theorem shows that no purely rational system can be proved just on the basis of its own precepts. We need to bring in some non-rational premise, for example from observation or basic philosophical sense; and, of course, to make and then test predictions derived from that premise. I am assuming that the mathematicians I am describing in this paragraph based their system on some empirically or intuitively obvious axiom.

12.1.5. Origin of religion

Finally, I should make an important comment: the arguments that I am presenting in this chapter are not the centre of Christianity or any other revealed religion. The apostles, as far as I know, had no or very limited knowledge of the philosophical arguments of God found in the Greek philosophers such as Plato and Aristotle: certainly there is no mention of arguments of this kind in the New Testament beyond an allusion to its possibility in the opening chapter of Paul's epistle to the Romans and a few hints in John's writings. The first attempt I am aware of to apply philosophy to Christianity came about a century after Christ (and little of Justin Martyr's work has survived to this day). To my knowledge, the Koran and Hadith contain no philosophical awareness at all; while Moses, of course, pre-dated Greek philosophy so could not have had access to it.

It is somewhat common (if you will excuse my somewhat simplified summary) to say that belief in God started when primitive man was left in awe by the wonders of nature, thunder, rivers, mountains, stars, rain and snowfall, death, the cycles of life and so on; so to explain such things started to anthropomorphise them: they seemed like elementary powers beyond his comprehension. As he began to construct and use tools, he saw an analogy with the world as the work of a toolmaker; as he gradually organised himself into societies and created informal laws wondered if there was not an even greater lawmaker; as the first of them began to seize power wondered if (and made use of it for their own ends to justify their authority and to subjugate their underlings) there was not a greater power; as they began to mould certain ceremonies to frame the changes of the seasons and sacrifices to offer atonement pondered if there was something some greater symbolism behind them; they searched for the meaning that life appears to have and found it in their own stories. Perhaps on occasions a few sacrifices by chance 'worked': they appeared to bring rain, or withhold the pestilence or locusts or brought victory in battle and these anecdotal events grew in the telling and gained a life of their own. Idle fireside stories and speculations were repeated, the stories became myths, the myths became religions, and the religions were seized by the political powers to justify themselves and establish control.

I cannot say that this process, or something like it, has never occurred. Certainly, the origins of religious belief occurred in pre-history, so we cannot tell if religion arose in this way or if it arose from an actual experience with God or a demon – there is simply no direct evidence available to us. But one thing is clear: with the possible exception of Hinduism (which does have its roots in pre-history), Buddhism (which, as far as I know, did originate from one man's contemplations) and the various primitive animists and spiritualist religions which still endure in certain parts of the world, none of the major religions around today claimed to have developed in this way, or are consistent with this picture. And for those founded in the historical period – after the invention of writing – it is recorded how they began. Every religion of any worth began in response to a supernatural event (or a supposedly supernatural event), which, if true, left the original witnesses in no doubt about the meaning of the message. Sometimes this occurred in visions; sometimes something more dramatic. A vision, of course, is a weaker argument than a physical sign – unless it is explaining something which could not be known, such as future events.

Furthermore, for the main monotheistic religions, this is not an event that occurred in the distant past at a nondescript time, such as the Greek myths, but set in a real historical environment. It is not something that was suddenly invented centuries later such as the Arthurian legends, but something fully embedded within culture from the beginning. An attempt to explain a religion must also explain these origin stories or myths. The philosophies and anthropomorphising did not come first, and the stories invented later; but the events came first, and the various dogmas and philosophies developed in response to the events. This can be seen from the development of any of the major religions; remove the various theories of atonement and salvation from Christianity, and you still have the accounts of the crucifixion and resurrection of Jesus; remove the crucifixion and resurrection

and nothing of Christianity remains at all, because all the dogma and doctrine assumes them. The event, therefore, had to have come first; the doctrine either accompanied it or followed it.

The philosophical arguments for God; the nature arguments for God; the 'warm fuzzy feeling' arguments for God all came later, as an attempt to either explain or understand the observations. This means that a philosophical argument for God is not enough to lead one to religion; equally attacking such arguments ought not to be enough to lead one away. If observation tells you one thing, and reasoned argument another, observation must always win: you know that the argument is flawed, even if you cannot see the flaw. Experience is crucial - but experience also is not enough by itself, at least not enough for a thinking man. If one sees a great wonder that convinces you of God, then one still faces Hume's argument against miracles. The empirical observation would be enough to convince you that Hume got it wrong, and there was a flaw in his reasoning or one of his premises. But to convince others, you still need to see what the flaw is, and be able to argue against it. Understanding or experience? We need both, especially since our understanding demands that we experience, and our experiences compel us to seek understanding.

The purpose of the arguments, then, is not so much to prove but to explain, and to provide a framework to answer criticism. This is no different from scientific investigation; scientific theory is not constructed to prove experiment but to understand it. Of course, the theory is the target of the whole endeavour; knowledge is the goal – but we get there through observation. Theory without experiment is groundless; experiment without accurate theory is pointless. So it is in Christianity. The goal of the Christian life is to know God and to be and to desire to be what God designed for us; and we get there by reason built on observation, all the while fully relying on God to shape and perfect our intellect, senses and character.

But before we can know God, we must first of all sweep away the prejudices that bind us and prevent us from seeing Him; and that is what philosophy, reason, and an open and inquiring mind can do.

12.1.6. Definition of God

I should also comment on the definition of God, because this too is a point of separation between enlightenment thinkers and classical theism. In classical philosophy, God is defined as a metaphysical first principle, the being that cannot be explained in terms of anything external to Itself (there are, of course, different but equivalent ways of expressing this which one would find in the various writers and which depend on the particular metaphysical assumptions of those writers; my own definition, which I will give later during my own argument, is one of these). The orthodox expressions of Christianity are based on classical philosophy, whether neo-Platonic or Aristotelian, and naturally adopted this definition. This is fine; if the God of Christianity exists then He is certainly a metaphysical first principle among other things. Of course, stating the definition does not show that such a first principle exists, nor that there is only one such being. Equally, although if the Christian God exists He satisfies this definition, it is not obvious that if there is a being which satisfies this principle then that being is the God identified by Christianity, or any other revealed religion. It must either be subsequently shown that the various attributes of the Christian God are consistent with this principle, or that they can be logically proved from the definition and perhaps some other premises.

Of course, this work has been done by the major classical Christian philosophers. This definition of God (or something equivalent to it) was accepted by the early Church, the Medieval Church, and by both sides in the reformation (baring those more radical protestants who dismissed all philosophy). It is also not specifically Christian; it is a common point of departure for Greek paganism, Judaism, Islam, and I believe similar concepts exist in the Eastern religions of Hinduism and Buddhism. The idea of a first principle is not even

inconsistent with some forms of atheism (note that modern atheists deny the existence of God as defined using expressions similar to those described in the next paragraphs), though other atheists would deny that there is a first principle. And this generality is good, because to have a discussion different viewpoints need a common reference point with which to start. Once we have established the existence of such a first principle, then we can go on to consider what attributes It must have.

However, the enlightenment philosophy, especially the empirical tradition, is incompatible with this definition, because it either views metaphysics in a different way, or denies metaphysics or states that the metaphysical is ultimately unknowable. An alternative definition was needed, and two were provided. These continue into the modern world; when people heavily influenced by modern philosophy think of God, which is almost everyone today in the West except classical Christians, Jews, Muslims, and those of other religious traditions, they think in terms of one of these two definitions. The first, which was inspired by Leibniz and Kant, was to use a superlative to provide the definition of God, saying that God was the greatest possible being, or the being most worthy of worship, or some other such thing. I dislike definitions of this type because in my view they tend to be either ill-defined or circular, and beg plenty of questions; I will expand on this dislike later.

The second modern definition of God is, in standard empiricist style, to define God in terms of His attributes. For example, to say that God is the immaterial, omnipotent, omnipresent, personal, omniscient etc. being. My objection to this is similar to my objection to this type of definition for any type of object, namely that it implicitly denies theoretical knowledge. Defining something in terms of its properties is always unsatisfactory, because it makes the properties, rather than the substance, more fundamental, when in reality the reverse is true; this approach is thus a hindrance to understanding the substance, which means that it is a hindrance to understanding. It should in principle be possible to deduce the attributes of God from knowledge of His essence; His attributes are in the order of understanding secondary, and therefore to use them as a basis for the definition is to get things the wrong way round. Of course, defining God in this way also allows for conceptions of God which are not metaphysical ultimates, more akin to Zeus or Thor, who are basically just extrapolated versions of ourselves, than the God of classical theism, who is an entirely different kind of being.

When Aquinas argues for God, it should always be borne in mind that he uses a variant of the classical definition of God. If one tries to put modern definitions into the mouth of an ancient author, then one will invariably come to a misunderstanding. This is one of the problems when a modern reader studies the opening pages of the Summa Theologica ripped out of the context of the rest of his writings and his underlying philosophy.

12.2. Classical Presentations of the Argument For God

Objection:
The argument used in the next chapter contains modified forms of arguments which have already been disproved by the early modern philosophers, such as Hume and Kant. Nobody now takes such arguments seriously. Therefore this whole thesis is a waste of time. •

Response:
First of all, I do not exactly reproduce the classical arguments. However, there are certainly enough similarities that if the classic 'refutations' of the arguments for God succeed, then my own approach will also have to be questioned. The problem with this objection is that it assumes that the early modern philosophers succeeded. That assumption needs to be questioned. One objection is that the modern philosophers adopted a metaphysics and various premises, which, though historically useful in the development of classical mechanics, have

since been shown to be false. If their premises are false, we should certainly be hesitant in saying that their conclusions are certain. But most pertinently, it is clear that the early modern philosophers (or at least those most commonly cited) did not answer the classical arguments for God which most closely resemble those arguments I use here because they did not even address them. ○

This response obviously needs further expansion. I choose to study the arguments of Hume and Kant, because they are most commonly (at least in my experience) cited in having addressed the cosmological and teleological arguments for God, and because most (not all, but in my experience most) current objections to cosmological arguments I have encountered are basically just re-workings and refinements of theirs. I do not wish to dwell in full detail on this subject since I don't want to focus on other people's arguments, but rather present my own; I will discuss various objections which are pertinent to my own arguments following the case I put for them. However, I think that some mention of this topic is required to answer various common misconceptions about what Hume and Kant showed. Again, to be brief, I will only discuss one medieval presentation of the cosmological and teleological arguments. I have decided to pit the modern philosophers against Thomas Aquinas rather than (say) the Islamic scholars or other Christian scholastics, partly because Aquinas is often cited as the main representative of medieval arguments for God, partly because I am most familiar with his work, and partly because his approach has the closest similarity to my own. My strategy will be to present and explain Aquinas' first and fifth ways; I will then briefly discuss Anselm's ontological argument, not because I intend to use or defend it, but because it is important in Kant's attack on the cosmological argument; finally I will present and critique Hume's account against the design and cosmological arguments as presented in his *Dialogues Concerning Natural Religion*, and Kant's argument against the cosmological and ontological arguments as presented in his *Critique of Pure Reason*. For a more comprehensive refutation of the various responses to the arguments for the existence of God, see (Feser, 2017).

12.2.1. Aquinas

Thomas Aquinas was a thirteenth century Roman Catholic theologian, originally from Italy, but whose most important work was completed in Paris and Germany. He lived at a pivotal time in Western European civilisation. Western Europe had, by now, firmly started to regain its prestige, which had been lost with the fifth century fall of Rome. Advances in agriculture helped boost the economies; militarily, as the crusades showed, Western Europe could hold its own against the foremost power of the time, the Islamic Caliphate. An intellectual society, based around Plato, had been established, and the early Church schools had morphed into the early universities. A rigorous curricula was founded, and with the common language of Latin, and the intellectual but not political sponsorship of the Church (which was common across Europe), a stable intellectual elite was formed, able (in principle, if not always in practice) to reach out via the Church's representatives in every town and village to the most able students. A continent wide intellectual community emerged, stable despite the numerous political and economic crises and hardships of the period.

This community was originally built around the works of Plato, Boethius and Augustine. However, at the end of the twelfth century, Latin translations of Aristotle and his commentators from the Islamic world (which, while led by a Muslim elite, still at this time contained a substantial Christian and Jewish population and intellectual class) began to circulate the continent. It is not clear whether these came back from the Muslims via the crusades or the reconquest of Spain, or from the Byzantine Empire (the Eastern Orthodox Christians) from their interactions with the crusaders, Byzantine colonies in Italy, or the conquest of former Byzantine territory in Sicily; but what is clear is that they caused a huge shock. Some denounced Aristotle's system as being inconsistent with Christianity; others,

perhaps most prominently Albert the Great, said that God had nothing to fear from rational inquiry, and that there was much in Aristotle and the commentators who had improved on his work that was good and useful. While Aristotle could not be completely accepted, with minor modifications the thoughts of the Macedonian could be made consistent with Christian revelation.

Aquinas was a pupil of Albert the Great, and he took from his mentor a great respect of the Aristotelian tradition. He was a formidable philosopher in his own right, one of the greatest there has ever been (although he did not see himself as such); his commentaries on Aristotle are unmatched to this day, and his original works such as *De Ente et Essentia* are brilliant in their own right. He was not satisfied to ape the Philosopher or the later writers of the Aristotelian tradition, but developed, corrected, improved and in many ways perfected the Aristotelian thought. He is, admittedly, probably more highly regarded in the past few centuries than he was in his own time and the period just after that. Then, he was seen as just one of a number of brilliant scholars. Now, he is often treated as though he were the only great medieval theologian, overshadowing the rival systems of Scotus, Ockham and others. Although Aquinas has always been revered, certainly in portions of the Roman Catholic Church, I think it is fair to say that in the generations immediately following him Scotus and Ockham had more prominence. Certainly the enlightenment philosopher's took a great deal more from those writers than they did from Aquinas (for example, Kant owed a great deal to Ockham). However, in the nineteenth century, there was a revival of Thomism in the Roman Church, as the papacy searched for a system to challenge the liberal philosophies without abandoning reason, and in the latter part of the twentieth century a number of secular philosophers either adopted Thomism or independently started to develop a very similar system as an answer to the Enlightenment's problems. In my view, the modern veneration is well merited.

But he is best known for his fusion of Aristotle, the neo-Platonism of Augustine, and the Roman Catholic interpretation of the early Jewish-Christian traditions. This was presented in two works: *Summa Contra Gentiles* and *Summa Theologicae*. The *Summa Theologicae* (the whole of theology) was a major, and unfortunately unfinished work (the form we have today was completed by Aquinas' students after his death).

The *Summa* is set up in a simple but effective question and answer format (each question is known as an article; confusingly the articles are arranged into more general topics known as questions). Aquinas raises a question, carefully posed so that the answer is either 'yes' or 'no' (at least most of the time; in a few cases he says 'yes' if you mean this by the question and 'no' if you mean that, based on different definitions given to various words by different authorities); raises some common objections to his answer, presents a brief response from some older authority (usually the Bible or some ancient Christian writer), then his own longer and often original answer, before replying to the objections. The whole system of thought then gradually builds up in a logical fashion, with answers from previous questions forming the premises to the conclusions he would derive later. It is an easy, well presented, and powerful format, which I wish would have been adopted by modern philosophers: it would have made them much easier to read, and much easier to follow their arguments[8].

The five ways are Aquinas' answer to the question 'Whether God exists?', and present five arguments to suggest that God does exist. This is the thirteenth article of the Summa Theologica. The first ten relate to the nature of theology and whether it is worth studying (not surprisingly, Aquinas answers 'yes' to this question: he argues that it is possible to know at least something about God via an *a posteriori* argument, from effects to causes. This is, of course, what he the proceeds to do: all the five ways start with simple observations about the world and work backwards to God). Then there are two questions on whether it

[8]Although obviously the Summa format is not the only suitable one; for example modern mathematicians with their clear enumeration of theorems, propositions and lemmas, each building on the others, have done equally well.

is possible to demonstrate the existence of God, and then we come to this issue. The first three of the five ways are variations on the cosmological argument, which dates back at least as far as Aristotle[9]: the arguments from motion (which was taken from Aristotle's physics), efficient causality, and existence. The fourth way is more difficult to categorise in modern terminology; it is perhaps most similar to a moral argument, but it is not really the same thing as a modern take on the argument from morality. Its origin lies in an argument for God made by Plato and his disciples, albeit considerably modernised by Aquinas. The fifth way, as far as I know original to Aquinas, is the teleological argument, or the argument from final causality (this should not be, but often is, confused with the argument from design). I will discuss the first and fifth ways here; my main argument (below) will be closest to a combination of the second and third ways (so to discuss those here would be redundant), and I make no use of an argument similar to the fourth way so I don't intend to discuss it further.

Before giving the arguments, there are a few things that need to be pointed out. Firstly, the arguments presented in the *Summa Theologica* are necessarily brief; Aquinas' intent was to give a overview of all of Roman Catholic doctrine, and that meant that he could only spend a limited amount of space on each topic. Most of the arguments of the *Summa Theologica* are more fleshed out either elsewhere in Aquinas' writings or in some other writer he respected. Thus our reaction when an argument seems to be insufficiently detailed should not be to dismiss it, but to try to hunt down and find the longer version he is referring to.

Secondly, the *Summa* was intended to be an introductory textbook for theology students at Europe's universities. Now, to understand what this implies about the intended readership, we have to know something about the Medieval university curricula. The first readers of the *Summa* were already convinced Roman Catholics. The purpose of this book was not so much to persuade as to provide intellectual satisfaction and education concerning matters already accepted; and the question on God's existence provides the premises for the rest of Aquinas' work. In addition, students entered the medieval university at a younger age than we do today; in today's terms it would be midway through secondary school. They would already be literate, and reasonably well read. The first stage of a university education was a thorough grounding in the numerous subjects seen as fundamental: grammar, mathematics (at that time little more than arithmetic, geometry and algebra), physics (including astronomy), rhetoric, and various branches of what we now regard as philosophy: logic, metaphysics and ethics. It was only after this that students studied one of the more advanced topics, such as theology, medicine or law. Alongside the more general overview offered by works such as the *Summa Theologica*, the theology students would also be given a thorough grounding in the Bible and ancient Christian writers, in part through collections such as Lombard's sentences, and in part through extensive commentaries (Aquinas' own are a particular highlight of the medieval tradition, and among the best I have read, leaving many modern authors in shame), and in part through direct study of the primary materials. In other words, Aquinas could assume that his readers had a thorough grounding in classical philosophy. He did not, for example, need to defend the principle of causality, or define underlying concepts such as act and potency – or even God, because this had already been done.

This is a difficulty for the modern reader, because philosophy is not taught as a main subject, and where it is taught it is usually the enlightenment philosophers who dominate, and even the brief mentions of Plato and Aristotle are interpreted through an enlightenment lens. This means that we are largely ignorant of the prerequisites of Aquinas' work, and the precise and technical definitions he assumes, and where we are not wholly ignorant we are often confused by a entirely different set of principles and way of thinking. To understand Aquinas properly, we need to lay aside everything we have learnt from Descartes and his successors, and instead pick up the medieval textbooks on philosophy to learn from them.

[9]Aristotle, *Physics* VII, 4-6

Once we understand Aquinas, we can and should implant Descartes etc. back into our brain in order to critique the scholastic, or perhaps to give a voice to the Aquinas we now understand properly to let him critique Descartes and his successors.

Thirdly, the five ways are merely Aquinas' opening move. He builds on them. Many modern critics look at the five ways and say, in effect, 'So what?' All this demonstrates is is that the critics didn't bother to read Aquinas any further than article three of question two, because he amply demonstrates the consequences of his proofs in the rest of the first part of the *Summa*. And this is the problem we face: there is a lot of material which comes before the student got to this point, and a lot of material which follows after the five ways; the first of these is now unfashionable, the second barely studied at all except in theology courses. But in a modern philosophy course, the five ways are presented, stripped of this context, and the student with no real grounding in classical philosophy or logic is meant to make of them what they will; it is hardly surprising that they are often misunderstood.

Finally, as is obvious from what is said so far, Aquinas operated in a framework which was essentially a refined and improved form of Aristotle's. He knew nothing of Galileo, nothing of Newton, nothing of Maxwell, nothing of Einstein, and nothing of Dirac, Dyson and Feynman. Many of his examples were illustrated with Aristotelian physics, and Aristotelian physics is wrong. The question is how much this matters. Contemporary defenders of Aquinas say that it is not so important: the arguments themselves are based on metaphysics; the illustrations can easily be changed and updated. On the contrary, the basic principles he requires: act, potency, form, matter and the four-fold causality are necessary presuppositions for any scientific investigation. My own feeling is that this cannot be presumed, but ought to be demonstrated; which is why I have spend some time discussing how much of this metaphysics survives quantum field theory in the previous chapters, and how much it needs to be modified.

It may be objected that I, too, am just stripping Aquinas out of context. To a certain extent, this objection is justified – I can only plead necessity, since this is not intended to be a book about Aquinas: I have discussed some of the philosophical prerequisites in the previous chapters, and I will later continue the discussion in a way that bears some resemblance to Aquinas' approach; but I am not fully consistent with Aquinas in my discussion of either the premises or conclusions to this article. I can only direct the reader to the relevant literature[10].

So with a rather long introduction and caveat out of the way, let us dive into the first way, which is based on Aristotle's own argument for God[11],

> It is certain, and evident to our senses, that in the world some things are in motion. Now whatever is in motion is put in motion by another, for nothing can be in motion except it is in potentiality to that towards which it is in motion; whereas a thing moves inasmuch as it is in act. For motion is nothing else than the reduction of something from potentiality to actuality. But nothing can be reduced from potentiality to actuality, except by something in a state of actuality. Thus that which is actually hot, as fire, makes wood, which is potentially hot, to be actually hot, and

[10]Most contemporary introductions to Aquinas will cover the basic background material better than I have here. I would particularly recommend at a more popular level (Feser, 2009; Kerr, 2009; Kreeft, 1990; Feser, 2014; Chesterton, 1933) and at a more scholarly level (Gilson, 1994, 2003; Maritain, 1962), although there are, of course, many good texts available. Of course, once one has mastered the background with a contemporary introduction, there is no better way to proceed than to read through the original material: and I would highly recommend Aquinas' commentaries on Aristotle's *Physics* and *Metaphysics*, his *De Ente et Essentia* and the opening parts of the *Summa Theologica* and the *Summa Contra Gentiles*.

[11]I here only present a small expansion of the argument; because of lack of space my presentation is intended to only slightly less of a summary than Aquinas'. There are many books written giving the full background and scope of this argument and its consequences; the best modern treatment I know of is (Augros, 2015).

> thereby moves and changes it. Now it is not possible that the same thing should be at once in actuality and potentiality in the same respect, but only in different respects. For what is actually hot cannot simultaneously be potentially hot; but it is simultaneously potentially cold. It is therefore impossible that in the same respect and in the same way a thing should be both mover and moved, i.e. that it should move itself. Therefore, whatever is in motion must be put in motion by another. If that by which it is put in motion be itself put in motion, then this also must needs be put in motion by another, and that by another again. But this cannot go on to infinity, because then there would be no first mover, and, consequently, no other mover; seeing that subsequent movers move only inasmuch as they are put in motion by the first mover; as the staff moves only because it is put in motion by the hand. Therefore it is necessary to arrive at a first mover, put in motion by no other; and this everyone understands to be God[12].

This requires some unpacking and explanation; I have used other writings by Aquinas, Aristotle, and the commentators of Aquinas to provide extra details not present in this brief text.

(1) *It is certain, and evident to our senses, that in the world some things are in motion.* Motion, of course, meant a different thing to Aquinas than it does to us. We need to think in terms of act and potency. So suppose we have some form x, which has potentia $|x_0\rangle, |x_1\rangle, |x_2\rangle$ and $|x_3\rangle$, and suppose that initially x was actually $|x_1\rangle$ and therefore potentially $|x_0\rangle$, $|x_2\rangle$ and $|x_3\rangle$ (I will treat $|x_0\rangle$ as non-existence, so we have the possibility of x coming in and out of existence). Motion means a change from one state to another, so $|x_1\rangle$ ceases to be actual and (for example) $|x_2\rangle$ becomes actual. In Aristotelian physics, this might mean a change in place (which is where our modern understanding of motion comes from); the equivalent thing in Newtonian physics would be a change in momentum. The word 'motion' is thus confusing, and I would drop it if it we were not stuck with the terminology for historical reasons and if there was a good alternative; one alternative which will appeal more to the Newtonian reader is to substitute 'acceleration' for motion, since acceleration plays a similar role in Newtonian physics as locomotion does in Aristotelian physics. However, this fails to account for the other forms of change in Aquinas' concept and focuses us too much on physics rather than the underlying metaphysics (a mistake made too frequently as it is). Many people suggest that we should just substitute 'change' wherever Aquinas wrote motion, but this too is not enough, because we have to understand change in terms of actualisation of potentia, and most modern people don't think in that way.

(2) *Now whatever is in motion is put in motion by another.* Aquinas subsequently argues for this statement.

(3) *Nothing can be in motion except it is in potentiality to that towards which it is in motion; whereas a thing moves inasmuch as it is in act. For motion is nothing else than the reduction of something from potentiality to actuality.* In my example, $|x_2\rangle$ is 'that towards which it is in motion' and $|x_1\rangle$ is the 'thing that moves'. The opening part of this phrase is defining $|x_2\rangle$ as a potentia of x, while the closing line is the definition of motion I summarised above.

(4) *But nothing can be reduced from potentiality to actuality, except by something in a state of actuality.* This is the fundamental principle of causality: only something which exists, i.e. is in a state of actuality, can initiate the change from $|x_1\rangle$ to $|x_2\rangle$. Note that this is not saying that only something in a state $|x_2\rangle$ can cause $|x_1\rangle$ to become $|x_2\rangle$; Aquinas' example of fire causing something to be hot may be a little

[12]Thomas Aquinas, *Summa Theologica*, I. II.3.

misleading to the modern reader. All Aquinas is saying is that whatever causes $|x_1\rangle$ to become $|x_2\rangle$ must be actual, whether that is $|x_1\rangle$ itself or something else such as y. Recall that when Aquinas discusses causality, he means a substance rather than an event cause.

(5) *Now it is not possible that the same thing should be at once in actuality and potentiality in the same respect, but only in different respects.* This should be obvious; x cannot simultaneously be actually $|x_1\rangle$ and actually $|x_2\rangle$ it must be one or the other or neither. Aquinas goes a bit further, and discusses the respect, i.e. a part of a body or a property. This forestalls the objection that one might have one part of a body hot and another part cold simultaneously, for example one end of a string might be on fire while the other end is bathed in ice. This is not, however, the same thing as saying that the string is simultaneously *in the same respect* hot and cold.

(6) *It is therefore impossible that in the same respect and in the same way a thing should be both mover and moved, i.e. that it should move itself.* In other words $|x_2\rangle$ (the moved) cannot initiate (as the mover) the change from being $|x_1\rangle$ to becoming $|x_2\rangle$, because $|x_2\rangle$ doesn't exist when the change starts. This should be obvious. It also might seem as though Aquinas is implying here that $|x_1\rangle$ can't be responsible for this change, but it must be something else, y (which, of course, must be actual itself). A reason for this could be that the potentia are meant to be stable states they don't change in time unless acted upon by something else. $|x_1\rangle$ naturally should just remain as $|x_1\rangle$, unless something (the mover) causes it to become $|x_2\rangle$. However, Aristotle's physics allows for natural motion in which an object in one potentia has within itself a natural tendency towards another potentia – an object (supposedly) falls because the element of earth has a natural tendency towards the centre of the universe – so my feeling is that Aquinas's argument was not intended to rule out the case where $|x_1\rangle$ contains within itself the reason for its change to $|x_2\rangle$. The point is that before $|x_2\rangle$, there has to be something, whether that is $|x_1\rangle$ alone or a combination of $|x_1\rangle$ and y.

(7) *If that by which it is put in motion be itself put in motion, then this also must needs be put in motion by another, and that by another again.* That initial 'If' is of huge importance. We now consider y. Let us define the state y is currently in as $|y_1\rangle$. There are several options: firstly, that y has only one potentia, so y can only be $|y_1\rangle$; this condition is known as pure actuality. In this case, y was not put into motion, since it can only ever be in the state that it is in: there is nothing it can even in principle change into. Secondly, y has several potentia, $|y_0\rangle, |y_1\rangle, |y_2\rangle, |y_3\rangle, \ldots$, but it happens to have remained in $|y_1\rangle$ indefinitely. Thirdly, y has several potentia, and it changed (say) from $|y_2\rangle$ to $|y_1\rangle$. In this third case, we can say that there must be something else, say z (or possibly x again) which provoked this change. This obviously just returns us to where we were: we can say that if z was put into motion, then something put it into motion. The second option does not seem feasible: to have several potentia means that it can possibly be in another state than what it is currently in; it can possibly have moved in the past. But if it can possibly have moved, then over an infinite period then it would have moved (that's ultimately what possibly means: given enough opportunity over a wide enough range of circumstances, it will happen), thus the only way for something to be not have been put in motion is that it is in a state of pure actuality. Otherwise, we have a chain of one thing being moved by another, which was moved by another, and so on.

(8) *But this cannot go on to infinity, because then there would be no first mover, and, consequently, no other mover; seeing that subsequent movers move only inasmuch as they are put in motion by the first mover; as the staff moves only because it is put*

in motion by the hand. Some interpret this argument to be circular, saying 'there cannot be no first mover because then there would be no first mover', but this is missing the point. In particular, the phrase 'first mover' is used in two different senses in this sentence; one time meaning as the first in a series, and the second as the originator of an instrumental series (which need not necessarily be first in the sequence of causes). We must distinguish between two types of series; the first (known as an essentially ordered series; I will sometimes call it an instrumental series) can be illustrated by the following example. We have some engine fixed to the ground next to some railway tracks. The engine is connected to a large flywheel, the flywheel is connected to a chain, and the chain runs along the tracks to a carriage. The rear of the carriage (which is incapable of self-locomotion) may either be connected to nothing, or to another carriage, and that carriage to a further carriage, and so on. It is clear that the carriages can only move if they are pulled along by the engine (or something that plays the same role as the engine)[13]. Now suppose that we come across the last carriage, and it is in locomotion; we do not directly observe the engine. There are two possibilities: either it is pulled along by another carriage, or directly by the engine. That other carriage might be pulled along by another carriage, or by the engine. Now, it is clear that if any carriage is in motion, there must be an engine (or something playing the same role as the engine). But suppose that someone objects to this, and says 'But what if there is an infinite number of carriages? Do we have to have an engine then?' The question is nonsensical: however many carriages we have, if there is motion, then there must be at least one engine, and this applies to however many carriages we have, including an infinite number (were that possible). There must be N carriages and an engine, where N is some unknown number. Even if we take the limit $N \to \infty$, there must still be an engine, so the infinite number of carriages doesn't evade the conclusion. Don't forget, in mathematics – aside from some uses in set theory analysis which aren't relevant for this discussion – and certainly in physics the concept of infinity is only defined in terms of a limit. We cannot start with infinity, but only move towards it as we change another variable. $1/0$ is undefined. $\lim_{\delta \to 0} 1/\delta$ does make more sense, since each step along the way is an actual physical number, and it is this quantity we equate with infinity. Equally, 'What if the carriages are arranged in a circle?' fails (mostly for the same reason, but also partly because we started by finding an end carriage, which rules out a circle).

But then there is a second type of series (an accidentally ordered series), which perhaps can be illustrated by us finding the rear of a carriage which is not in locomotion. The front of the carriage is either coupled to another carriage, an engine, or nothing. In this case, there is no good metaphysical reason why there should not be an infinite number of carriages and no engine (there are good practical reasons, of course, but they arise from the limitations of the analogy rather than the underlying metaphysics). Now, the engine represents the state of pure actuality; the last carriage x, the penultimate carriage y, and so on. The state of locomotion represents a change from $|x_1\rangle$ to $|x_2\rangle$. The carriages, like x, y, z and so on are not capable of self-motion.

The main way that we can distinguish between an essentially ordered series and accidentally ordered series is that in an essentially ordered series every member

[13]Before someone, possibly named Galileo, starts quibbling about frictionless tracks, and no air resistance, and then invokes the principle of inertia, I will just say this: if you wish to change the picture in this way, then please also feel free to change every mention of 'locomotion' in this argument with 'acceleration' and 'at rest' with 'constant velocity'. You will end up with the same conclusion. In the real world, however, – or at least as close to the real world as we can get while postulating an infinite number of carriages – we have to deal with friction and 'locomotion' is appropriate.

of the series is in a state which depends on its neighbour also having been in a particular state at the key interaction, and its neighbour's neighbour, and so on down the line. Physically, we would say that some of the states are excited (for simple particles) or away from thermodynamic equilibrium (for systems of particles). If one carriage isn't being pulled, then the next one won't be pulled either, and neither will the one a hundred carriages down the line; or indeed any of the carriages before or after our particular sample (recall that the first way is the argument from motion rather than from causality; technically it is more concerned with the power to generate or be affected by a force than it is with efficient causality in itself). It is a series of instrumental causes. However, in an accidentally ordered series, a carriage only cares about the two on either side of it. It is not affected by the presence, absence or state of the carriages beyond those. In an essentially ordered series, each carriage has to be in a specific state, which forces us to ask why it is in that state rather than another, while in an accidentally ordered series it could be in any state, so the question of why it is in that state is not relevant to the existence of the series: it has to be in some state, so since it could be in any state without affecting the series, why not that one?

I think that there are four main distinguishing features of an instrumental series:

- Something is carried down the line of causes. In both the case of the carriages, and Aristotle's example of the hand moving a stick which pushes various stones, this is the force that is applied from cause to effect. The engine pulls the first carriage by some mechanical force, which consequently means that a force is applied by the first carriage which pulls the second carriage, and so on. The power to generate this force is passed on down the chain. This thing that is passed on need not be identical throughout the series; it can be converted from one form to another. The stored chemical energy that moves the muscles in the arm is different from the mechanical force that moves the stone.
- The causes cannot by themselves generate this thing that is carried down the chain of causation. The stone cannot push the next stone unless it itself has previously been pushed. A carriage cannot move unless it is first moved. A different type of being is required to give the first push. Thus the particular state of each being can't be explained without reference to the earlier members, and there is a dependency on the previous members of the chain.
- There is a direction in the sequence of causes. The stone pushes the next stone, which pushes the next stone, and so on. One of them is the pusher, the other the pushed. In terms of the same relationship, they do not both act as pushers to each other; neither do they both act as pullers to each other. Rather each stone is pushed by the previous one in the sequence and pushes the next one. In particular, this directionality is towards some equilibrium state. The thing that passes down the chain of causes naturally disperses. In both the case of the carriages and the stones, friction gradually ekes away at the motion until the objects are at rest with respect to the ground. If the sequence had continued forever, then the stones would be at rest; therefore there has to be something constantly providing them with sufficient force to overcome the friction. That something cannot be a stone or carriage, but must be a different type of being.
- There is direct contact between neighbours in the chain, but not necessarily across the whole chain. The act of stone A pushing is simultaneous with and local to the act of stone B being pushed. That doesn't mean that the whole motion along the chain is instantaneous; one can have a delay between B being

moved and B moving C. For example we can think of the energy being stored up for a few seconds within B, as energy is stored in a spring or the electric charge being stored up in a capacitor, before being released and subsequently acting on C. But B's moving C will be another local and simultaneous action; it just need not be simultaneous with A moving B. Without this contact, it would be impossible to attribute the motion of B to the act of A.

So the question reduces to which of these pictures most closely resembles the chain of movers and moved: the first option, where an infinite chain of carriages is impossible, or the second picture, where an infinite chain of carriages is possible? Aquinas then gives the example of staff being moved by the hand. This is an example of instrumental causality; in our picture we can say that the final carriage is pulled by the engine, with the intermediate carriages merely tools the engine uses to produce the effect. In the same way, we can say that x is moved by z, with y as an instrument used by z to move x. Now any instrumental series must have a first mover and a last mover; the first mover is necessary to provide the impetus for the whole chain of motion. If there were no first mover, there would be no fundamental reason why anything would be moving. Thus Aquinas wants to argue that the first picture, with the carriages in locomotion is closer to the metaphysics rather than the second picture with the stationary carriages. Of course, it doesn't matter that in nature there are different types of carriages (corresponding to the different kinds of fundamental particle), as long as they all lack the property of being able to move themselves (in the sense that the state it is moved to cannot generate the motion). We can, perhaps imagine a carriage with an inbuilt engine within it, but that still shows that there is something other than the fundamental particles we are familiar with from physics out there (which is, of course, in part the goal of the cosmological argument), and it is also a point in the argument where the analogy breaks down: the engine doesn't in reality move itself, but has various parts which are moved by the burning of some fuel, with the fuel moved by the various chemical reactions of their individual molecules which are moved by .... In short, this example just brings us back to the chain of causality described in the first way; however in the analogy the carriages stand in for the most fundamental physical objects, so such a case cannot happen since they can't be broken down into parts in the same way that a train engine can.

I should also note one argument presented by Aquinas in another work (I am not sufficiently knowledgeable about scholastic philosophy to know if this was the only or best argument). Aquinas held that because for most things existence and essence are separate, it is necessary to explain not only how a being reaches a certain state but also why it remains in a certain state. There are two ways (according to Aquinas) in which this can be explained: 1) A being remains in a particular state, because it cannot not be in that state; 2) It is sustained in that state by something else. The first of these only has one possible state, which is actual (it has no potentiality), so this is a being of pure actuality, and a being of pure actuality is Aquinas' definition of God. The second possibility leads us to a regress, which can only be terminated by a being in the first category or continue forever. However, this series is manifestly essentially ordered, since everything remaining in that state requires that there is ultimately a being of pure actuality to explain anything (it is a series of instrumental causes, acting in some sense 'simultaneously').

Can this idea stand in the light of modern physics? One response is to say that physical law explains why things continue as they are, but this is unsatisfactory. The laws of physics are descriptive rather than prescriptive. In other words, to invoke (say) some principle of inertia to try to explain why it continues in that state

is not an explanation, but just a rephrasing of the question: the principle of inertia means that 'things tend to stay in the state they are in', which is nothing but a description of the problem we are trying to resolve. It doesn't answer the problem, only rephrases it in more technical language. Obviously a more sophisticated physicist would relate the principle of inertia to conservation laws, and conservation laws to symmetries, but this again fails as an explanation. Firstly, the link between conservation laws and symmetries assumes the principle of least action, which needs to be justified from first principles. Secondly, we can ask why are those symmetries satisfied? Why does an electron care about symmetry under translation? In other words, such a regress again only rephrases the question, rather than answering it. And thirdly, such an explanation doesn't help anyway, because the conservation of momentum is not a sufficient explanation for the principle of inertia: a body can change its state through an interaction with something else without violating the conservation laws.

Another explanation for continued existence might be that 'It just does.' This is intellectually unsatisfying, and at best only valid in a classical system where there is no creation or destruction of matter. For quantum field theory teaches us that there are occasions when 'It just doesn't': when an electron spontaneously emits a virtual photon, for example, and moves from one state to another. We have a situation where there is stability for short time periods (enough to require an explanation), but things are not permanently stable. We cannot just say that it is something's nature to continue in a certain state, because it isn't always. Existence and essence are not co-joined for anything except a being of pure actuality. Perhaps somebody might argue that there is no reason; but this would not explain why the decay rates are predictable and different for each type of being: if there is truly 'no reason' then we have a system which is irrational, where anything could happen, which is not what we see in physics. Nor can we say that there is some reason, but the explanation is unknown, and beyond what Aquinas could have understood. Aquinas' argument is general. It does not go into the microscopic details (which will be buried somewhere in physics well beyond what was available in Aquinas's day, and possibly even our own day), but claims that any explanation that does actually explain must satisfy the basic underlying principles. Whatever the true explanation is, this argument would apply. The precise details do not affect it.

A further objection might be that Aquinas fails to appreciate the distinction between 'is sustained' and 'is being sustained'; in other words the principle which causes the sustaining might not be a being as such; and if so his attempt at a causal explanation might be foolhardy. It might, perhaps, be some abstract principle that ensures that the body continues in the state it is in. However, I have defined existence as the capability to interact with something else; a sustaining abstract principle (if it were not a description of the relationship between two different types of being) would interact with something else; and so it would exist, and if it exists it must be a type of being, though possibly very different type of being from what we are used to (it would, for example, be entirely immaterial). Thus I don't think that this allows us to evade Aquinas' premise. This argument might not be obviously convincing to the modern reader, but I don't think that it can be dismissed as easily as people might intuitively think.

(9) *Therefore it is necessary to arrive at a first mover, put in motion by no other; and this everyone understands to be God.* Thus we come to the notion of the unmoved mover (which is better expressed as the unmovable mover), and the point of the argument. Aquinas is arguing because some things are in motion, there must therefore be at least one thing which is necessarily not in motion, which nonetheless has the ability to move everything else (either directly or instrumentally). Equally,

in the second way, Aquinas argues that because some things are caused, there must be at least one thing which is necessarily uncaused, which nonetheless is the cause of everything else. In the third way, Aquinas argues that because there are some things which can come in and out of existence, there must be at least one thing which must always remain in existence, because of which everything else can become existent. In all these examples, we are pointed towards a being of pure actuality as the only possible terminus of the chain.

Does everyone understand this to be God? Perhaps Aquinas should have said 'Everyone who has some understanding of Classical Philosophy or the Abrahamic religions understands this to be God,' which in his day included, of course, everyone on his side of India and the Sahara. Today, unfortunately, there is a lot of ignorance, and many people define God to be something a bit like us only with greater power and without a body. In the Bible, however, it is clear that God is the uncreated creator; He may be more than that, but He is certainly not less than that. Indeed, this is the classical definition of God. An uncreated creator is also obviously an unmoved mover, since creation is a form of motion (in the Aristotelian definition). Thus the unmovable mover is understood by everyone who is not ignorant to be God.

Note that Aquinas' argument does not assume that the universe had a beginning; it is equally at home in an eternal universe (or multi-verse) or a finite one. Aquinas just argues that there cannot be an infinite series of instrumental or secondary causes; a finite series of causes can, of course, take an arbitrarily long time, and we can take the limit towards an infinite amount of time between causes; equally the first cause was later in Aquinas' work shown to be timeless and eternal, so there is no reason why there should not be several series of secondary causes all going back to the same first cause, but beginning at different times. Indeed, Aquinas did not believe that it was possible to show that the universe had a beginning through reason alone. To assume that the universe had a beginning made life too easy for the theist, so Aquinas didn't.

So how do we evaluate this argument? There seem to me to be four possible points of attack. The first is to attack the reasoning that rules out an infinite chain of movers. The second is to look at the unmoved mover, and ask firstly whether it is indeed God, and secondly whether an unmovable being is capable of being a mover. The third is to question the underlying metaphysics. The fourth is to question the premise that things cannot move themselves. I do not wish to defend this argument in detail, not because I don't think it indefensible, but because my purpose is to present it to evaluate whether Hume and Kant's objections are sound; and I have already presented it in sufficient detail for that. I should also note that of these objections, only the first is even conceivably applicable to the second and third ways. The second objection Aquinas answered well enough himself in the *Summa Theologica*. Aristotelian causality links beings, not events, and the conservation of four momentum and positivity of energy means that if something is caused then it must have a cause other than itself; and this answers the equivalents of the third and fourth objections (which I consider to be the strongest) which could be applied to the second and third ways.

(1) I think the most common response to the first way is to challenge the idea that a series of movers is *essentially ordered* rather than *accidentally ordered*, or to ask whether a series of instrumental causes necessarily implies a first mover. Note that the question is not whether the series can extend infinitely backwards in time (Aquinas himself did not think that he could rule out an infinite past through rational arguments, but only through revelation), but whether it can extend backwards through an infinite series of movers none of which are capable of imparting motion unless they have been moved themselves (for example, moved from non-existence to existence, or from some actuality which is not capable of producing the observed effect to one that is capable of it). With regards to the analogy of the carriages,

the case of them being in locomotion is closer to the sequence of causes and effects. If the carriages are stationary, we only need explain their existence; if they are in locomotion, we need to explain both their existence and some property which they are not capable of generating themselves. With matter excited beyond the ground state or into the ground state (which is the modern day equivalent of Aquinas' system of movers), we not only need to explain the existence of such matter but its excitation, so once again there are additional properties. Matter is not capable of exciting itself. Similarly, if we start a chain of processes, we can consider an initial state and investigate the probability that a final state will arise. There are, however, a chain of intermediate steps between the two. These intermediate steps each represent the various movers and moved objects, but from the perspective of the series as a whole they are just instrumental causes linking the initial to the final state. Thus it is, I think, quite reasonable to claim that the sequence of actualisation of potentia as an instrumental, and thus essentially ordered, series.

For example, only something that has been excited can excite a being from the ground state (lowest energy state). A photon (for example) does not possess this ability in itself, but only because it has been given enough energy to do so. Similarly, an electron in some potential cannot emit a photon in itself, but only if it has been excited from the ground state. If we grant Aquinas' other premises, none of these fundamental particles are capable of moving unless they have been moved. The electron doesn't just require that the photon which excited it has the right energy, but whatever emitted that photon had the right energy levels, and whatever excited that, and so on. It doesn't just depend on the state of the photon being 'right' for its motion, but requires all members of the series going back either to the first member or to infinity to have the right properties. Thus the series is instrumental.

Does an instrumental series have to have a first cause? The causal power of each of the instruments is dependent on there being something with inherent causal power. They have derived whatever causal power they have from this immovable mover, and thus can only possess this power if there is an immovable mover at the start of the series. But since they do have the ability to actualise potentials, there must then be an immovable mover. Thus I think an instrumental series must always have a first mover. If the rest of Aquinas' argument is sound (in particular that self movement is impossible), then this point is as well.

(2) The second objection is to ask whether this immovable mover is indeed possible. After all, every action has an equal and opposite reaction. If it is capable of moving, it must be capable of being moved. This argument is, however flawed: even in the context of Newtonian physics, it is conceivable to think of a body with infinite mass. This would move (via gravity if nothing else) but could not be moved itself. But we need not postulate such things. Newton's third law only applies to physical objects, and the first mover is not physical: there is no reason why it should be bound by it. If the rest of Aquinas' argument is sound, we have shown that there must be an immovable mover, therefore the immovable being must be capable of causing motion in other beings. Secondly, one can question whether this is God. Now we should be careful here: by 'God' Aquinas (and Aristotle before him) does not necessarily mean the Christian God, with everything else that implies. Instead, we should ask what the most fundamental definition of the term means; and above all else God is understood as the uncreated creator. It is clearly not much of a stretch, given that creation is a form of motion, to reach our unmovable mover. Today people often define God as the uncreated creator plus other properties, but I think that that is mostly due to us having been brought up in a post-Christian society. Instead, we should think of what the term meant in the pre-Christian and

early Christian world, since it was these classical philosophers who formulated the language which was originally used to describe Christian doctrine. We are not discussing the numerous pagan gods, but the word 'God' in the singular refers to the primeval force from which even they originated. The Greeks, of course, did not regard this as a person; yet they still used the word 'God' to describe it. Thus, I think that use of the word 'God' is appropriate here. Certainly, the Christian God fits this definition, even if He is also considerably more than it. Thus again, if the third and fourth objections can be answered, I don't think that this is an issue.

(3) The third objection concerns whether the underlying metaphysics is accurate. Obviously, a Newtonian would have seriously questioned it, but, as we have seen, classical mechanics does not provide a sound base for questioning the metaphysics. Quantum physics seems to fit in well with the language of actuality and potency. But does it fit well enough? The system I formulated in previous chapters was not identical to Aristotle's. In particular, we have the worry that things can seemingly be in a linear supposition of states. For example, we may write a particle's wavefunction as $\alpha|x_1\rangle + \beta|x_2\rangle$, with α (initially) very close to one and β (initially) very close to zero. But β is not zero. Does this mean that $|x_2\rangle$ is in fact partially actual before the transition of the wavefunction from being dominated by $|x_1\rangle$ to being dominated by $|x_2\rangle$? And if so, does that mean that $|x_2\rangle$ can cause itself? We have to be careful here. The wavefunction does not describe the particle itself, but only our knowledge of it. The particle itself is either in a state $|x_1\rangle$ or $|x_2\rangle$ - whenever we measure it, we find one of those answers (or another potentia). The wavefunction merely means that the physics is not predictable (in the sense that it is not deterministic): that knowledge of the universe at one time does not allow us to predict what it would be at a subsequent time to perfect precision. It does not mean that the particle is itself confused about which state it is in at any given time. The wavefunction is a mathematical function but Aquinas' argument is not about mathematical functions, but what really is. Mathematics can be used to describe reality (or in this case a probability distribution of what reality might be), but it is not reality itself. Thus, I don't think that this objection definitely holds up; although I have to add the caveat that the interpretation of quantum mechanics is widely disputed, so I would need to do a lot more work than I intend to in this brief discussion to show that the interpretation I have used here is correct.

(4) Finally, is it true that beings cannot move themselves? Aquinas and Aristotle presented some seemingly metaphysical arguments to suggest that this is the case (in their other writings, the *Summa* itself doesn't go into this detail); but these are for nothing if physics says the opposite. This is the problem of spontaneous decay. The most common example of this that affects our day to day lives is radioactive decay, but I will instead consider the spontaneous emission of a photon by a free electron. The underlying principles are the same, but the mathematics somewhat easier. It is often said that there is no physical cause for the decay. So we start with an electron in state $|x_1\rangle$. A moment later we have an electron in state $|x_2\rangle$ and photon in state $|y_2\rangle$. Nothing interfered with the electron to encourage this decay – this is, of course, not true in the opposite case of the absorption of a photon, where we can say that the photon $|y_1\rangle$ is the mover of $|x_1\rangle$. But not in emission. No other particle $|y_1\rangle$ is involved. Does this destroy Aquinas' argument?

The first and most obvious answer is that it ultimately makes no difference to the argument. There is no reason why we have to introduce a $|y_1\rangle$ with each movement of x, since $|x_1\rangle$ itself has all the same properties of $|y_1\rangle$ that we need; namely, that it is actual and has some potencies. We do, of course, need to be able to go back from x to y at some point, otherwise we would be stuck circulating

around the various potencies of x and will never be able to reach our immovable mover. But that is allowed in physics as well. So, I think that in this way one can show that the objection is not a serious obstacle: it can be circumvented by a minor modification of the argument.

But I think that it is instructive to go into a little more detail. What is $|x_1\rangle$? We defined the potencies as the various eigenstates of the Hamiltonian. Which Hamiltonian? There are two we could consider: firstly the Hamiltonian of the free electron, H_x, and secondly the Hamiltonian of the entire universe, H_U, (with $|x_1\rangle$ being that part of the eigenstate that refers to the electron). Only the eigenstates of the universal Hamiltonian are stable, since this is the operator that evolves time, but this is no good for our present study, since we are just interested in the affairs a single particle, and the universal Hamiltonian describes everything. So $|x_1\rangle$ is the eigenstate of the free electron Hamiltonian. Now we know how x evolves in time: if it is in a state $|x_1\rangle$ at a particular time, then the wavefunction an infinitesimal time δt later is $e^{iH_U\delta t}|x_1\rangle \sim (1 + iH_U\delta_t)|x_1\rangle$. $|x_1\rangle$ is not an eigenstate of H_U, so $H_U|x_1\rangle$ contains $|x_2\rangle$. Thus there is a non-zero probability of spontaneous decay. In fact, we can be even more specific. if $\hat{x}_1^\dagger$ is the creation operator for $|x_1\rangle$, and $\hat{y}_2^\dagger$ the creation operator for the photon described by $|y_2\rangle$, then the term in the Hamiltonian which leads to the spontaneous decay we are considering is proportional to $\hat{x}_2^\dagger\hat{y}_2^\dagger\hat{x}_1$. And that is where the photon comes from – the universal Hamiltonian. In other words, the photon does not come from nothing, since the Hamiltonian is not nothing, but it is the action of the Hamiltonian on the state $|x_1\rangle$ that causes the emission. So $|x_1\rangle$ does not move itself, but is moved by whatever the Hamiltonian represents.

That last statement begs a rather obvious question, but I will leave that question aside for the moment, and consider H_U itself. Can the Hamiltonian itself be different than it is? Most physicists would answer that in some sense it can (as far as we know): it depends on a number of parameters that could in principle take any value. So we have a string of possible Hamiltonians, the one that describes our world, $H_U^{(1)}$, and then we have $H_U^{(2)}$, $H_U^{(3)}$ and so on which describe the time evolution of other possible universes. Once again, we have a picture of potentia emerging. $H_U^{(1)}$ is actual, but, in principle at least, it could potentially be $H_U^{(2)}$, $H_U^{(3)}$. So the question emerges is $H_U^{(1)}$ capable of being moved to one of these alternative Hamiltonians? This we don't know – it goes way beyond any physics – but logically there are two possible answers, yes and no. If no, then we seem to have Aquinas' unmoved mover (though in practice, we need a little bit more work to get there). If yes, then then since it is not capable of changing itself (since it is the operator that institutes change, and it does not act on itself) and we know from physics that no material being can change it, then whatever changes it must be some immaterial mover. We are now back on course with the first way, and heading towards God (assuming, of course that the metaphysical arguments against self-motion hold up).

But what of the other possibility, that H_U is not in practice capable of being changed? Does that make it the unmoved mover? Is the Hamiltonian itself God? No, because the Hamiltonian is just a mathematical function. As such, since a mathematical function is not itself a being, but only describes something, it cannot be God. Since it is an operator, it describes the interaction between two or more beings. We know that it interacts with matter, which leaves us with two possibilities. Firstly, it might describe the interaction between matter and some other lump of matter, or it might describe the relation between various pieces of matter and something immaterial. In the first case, the visible universe itself gives rise to the Hamiltonian; in the other, the visible universe plus something else is

> behind it. But we must recall that what is acted on by the Hamiltonian is not matter itself, but a probability amplitude. Now, probability expresses a gap in what is known for certain - if we say that something is probable then we mean that there are some missing ingredients in the available knowledge. However, this is not our own knowledge – the limits of experimental precision, for example - that is deficient: the wavefunction refers to a fundamental uncertainty in the material universe, something the universe does not 'know' (in a metaphorical sense) about itself. But every bit of information relating to the state of the material universe is contained within the universe; if the universe itself were sufficient then it would ultimately be deterministic rather than probabilistic. If the Hamiltonian only described the interaction between two particles in the universe, then it would surely act on some mathematical objects that describe the actual state of the particles themselves rather than a probability distribution for what those states might be. There would be no missing information which could give rise to a probability distribution. But the Hamiltonian does not act on the actual electron state, so therefore it cannot describe the interaction between the electron and something else in the material universe which is fully known within the context of the universe. This implies that the Hamiltonian does not just describe the relationship between material beings, but material beings and something outside the set of all material beings, which must therefore be some immaterial being. Now we are back on course with the first way, and quantum physics is no longer relevant - this being must either be in a state of pure actuality, or capable of change. If it is capable of being moved, then there must be some mover capable of moving it, and we can continue until we reach God.
>
> This is not to say that the metaphysical arguments against self-motion are correct. But, I think, it does show that the obvious objection from physics is not as strong as it may at first seem.

I therefore think that some form of the first way is still defensible, although the arguments of the thirteenth century require considerable modification and strengthening to be consistent with our modern understanding of physics. Obviously, I have not fully defended it here; it would take more space than I am prepared to give to it to adequately rephrase it and show that the objections fail. I do not think that that is necessary; the second and third way are not dependent on Aristotelian physics in the same way and are not challenged by modern physics, and they by themselves are good enough to get to where we need to go. Obviously, objections can be and have been raised against the second and third ways, but I will discuss any pertinent objections I am aware of later.

Now let's turn to the teleological argument.

> The fifth way is taken from the governance of the world. We see that things which lack intelligence, such as natural bodies, act for an end, and this is evident from their acting always, or nearly always, in the same way, so as to obtain the best result. Hence it is plain that not fortuitously, but designedly, do they achieve their end. Now whatever lacks intelligence cannot move towards an end, unless it be directed by some being endowed with knowledge and intelligence; as the arrow is shot to its mark by the archer. Therefore some intelligent being exists by whom all natural things are directed to their end; and this being we call God.

Again, it is useful to break this down and expand each step.

(1) *We see that things which lack intelligence, such as natural bodies.* Note well 'lack intelligence': this will be of key importance in a moment.

(2) *[These things] act for an end, and this is evident from their acting always, or nearly always, in the same way, so as to obtain the best result.* Thus Aquinas argues for

final causality from the regularity in the world. That A often leads to B but not C means that A is directed towards B but not directed towards C. I have already defended the concept of final causality, so need not do so again.

(3) *Hence it is plain that not fortuitously, but designedly, do they achieve their end.* That A leads to B is not caused by chance, because if it was caused by chance there would be occasions when $A \rightarrow B$ (where B here represents the entire set of final causes of A) did not happen. 'Chance' is not the cause of anything. That implies that there is something about A that encourages the outcome B (or encourages it to occur with particularly high frequency). The word 'designedly' means that this is a part of the nature of A.

(4) *Now whatever lacks intelligence cannot move towards an end, unless it be directed by some being endowed with knowledge and intelligence; as the arrow is shot to its mark by the archer.* This is the key premise. A lump of matter does not just say to itself one day 'I know, today I will become B, but I don't fancy becoming C.' There is no reason inherent in matter itself why it should be directed towards one end rather than another. It is thus something external to the matter which imposes this sense of directionality upon it. This external thing could, perhaps, be some other brute lump of matter, but then we have just regenerated the same problem at a different level. That other brute lump of matter must have directionality, since it directs A to become B but not C, and therefore it too needs to have that directionality explained. Ultimately (and again, one can rule out an infinite series of essentially ordered causes for the same reasons as before), the only thing (according to Aquinas) that can impose direction on something else is the will of some rational agent. Intelligences, of course, can say 'Today I will produce effect B rather than C'; we know this from personal experience.

(5) *Therefore some intelligent being exists by whom all natural things are directed to their end; and this being we call God.* The conclusion clearly follows from what proceeded it. Once again, one can question whether this being is God, but this being imparts directionality without having directionality imparted upon it. The final cause is the opposite of the efficient cause; if X imparts A with the disposition to become B, then X is at least in part the efficient cause of A, since it explains that aspect of As nature. But the final cause of the termination of the series of final causes is not imposed on it from outside, but must be intrinsic to its nature; therefore this being cannot have an efficient cause, and the only type of being which could exist but which cannot even in principle have an efficient cause is a being of pure actuality, i.e. God.

So what do we make of the key premise, that directionality can only come from an intelligence? Of course, the early modern philosophers tried to evade this question by denying the existence of final causality, but, as we have seen, this denial is not really tenable. One way to try to answer the issue is by saying that they are directed by the laws of nature, but this is an evasion rather than a response. It is not an answer; just a rephrasing of the question. It only brings us to ask, 'Why do things obey the laws of nature', and 'Why do the laws of nature direct them so?' One possibility is that the laws of nature reflect the will of some intelligence that directs natural bodies; and this is equivalent to Aquinas' answer (Aquinas, of course, as an Aristotelian, would not have thought about 'laws of physics' in the same way that we do; but if we try to translate what Aquinas did mean into a more familiar language we will be left with something similar to my phrasing here). Another response is to say that the laws of nature are some brute fact, so the question is ultimately unanswerable. It is curious, then, that Aquinas provided an answer to this supposedly unanswerable question. But in any case, the laws of nature are immaterial; I defined a sense of directionality in something immaterial as a will, so the laws of nature must reflect some will. Whether this will is exactly the same in its origin and nature as our will is not important (we don't

expect that to be true for God); nonetheless the directionality implies an immaterial will. Incidentally, Aquinas" argument against an infinite series seems to be more powerful in the context of the fifth way, if we grant the rest of his assumptions. If directionality ultimately requires an intelligence, then an infinite series of instrumental causes clearly cannot provide that directionality: there still has to be an intelligence.

Of course, to get fully to God as the uncreated wilful creator, and to enumerate the usual divine attributes, we need to combine the cosmological and teleological arguments, and provide further argumentation, which Aquinas did for the next several hundred pages of the *Summa Theologica*. As I mentioned, the five ways are only Aquinas' opening step, not the completion of his argument. But nonetheless, the arguments are there. I think the second to fifth ways certainly look reasonably sound (albeit that they might require some minor modification or enhancement in places to account for what we have learnt in the past eight centuries since they were first formulated). The first way perhaps requires some more major modification to reformulate it in a state consistent with quantum field theory, but this seems (to me) to be possible. So the argument has been made; the question asked. In a moment, we will see how well Hume and Kant answered it.

But it is important to again emphasise that Aquinas' arguments all rest on his metaphysics. From Newton to the early part of the twentieth century, an atheist who neglected the force of these arguments on account of his mechanistic world-view would have some justification. But, as we have seen, quantum field theory has drawn us away from the geometric approach to physics assumed by naive mechanism towards a more causality based approach. This brings Aquinas back into play; his arguments again need to be answered in his own terms.

12.2.2. The Ontological Argument.

First, however, I want to move a couple of centuries prior to Aquinas to Anselm of Canterbury. This is not an argument that I intend to use. Certainly, there have been reformulations of it more recently, most notably by Plantinga (Plantinga, 1965, 1979), which require consideration. However, I will need to discuss Kant's critique of this argument (or perhaps Descartes' or Leibniz's reformulation of it), partly because Kant used that critique as part of his answer to the cosmological argument, and partly because Kant's thoughts, if accepted, may be the basis of various objections against my own position, and I want to clear away those obstacles from the road before I start to travel down it. Therefore, before coming to Kant, I need to present the argument that he attacked.

> Thus even the fool is convinced that something than which nothing greater can be conceived is in the understanding, since when he hears this, he understands it; and whatever is understood is in the understanding. And certainly that than which a greater cannot be conceived cannot be in the understanding alone. For if it is even in the understanding alone, it can be conceived to exist in reality also, which is greater. Thus if that than which a greater cannot be conceived is in the understanding alone, then that than which a greater cannot be conceived is itself that than which a greater can be conceived. But surely this cannot be. Thus without doubt something than which a greater cannot be conceived exists, both in the understanding and in reality[14].

I have never really found this argument convincing. It seems to me that a robust definition of 'greatness' can only be determined once we know the type of being it refers to. For something to be great, it perfectly fulfils its nature. We thus need to know the nature of God before we can understand what it means for God to be great, and therefore using greatness as part of the definition of God makes the argument circular. One would need to

[14]Anselm, *Proslogion*, II

be sure that a consistent definition of greatness was used throughout the argument, which includes existence as part of what makes something great. Nonetheless, the ontological argument has frequently caught the eyes of many philosophers since Anselm's day.

12.3. Hume

So let us turn to Hume's response to the cosmological and teleological arguments. It is not clear that Hume had Aquinas in even a small corner of his mind when he wrote his *Dialogues*, and I am pretty certain that he didn't but rather was attempting to answer his contemporaries, but, nonetheless, since Hume is often cited as answering the cosmological argument in general, which includes Aquinas' version of it, it is still pertinent to compare the two sets of arguments. While I reject Hume's writings as they are mostly built on what are, to the one trained to think logically by Aristotle, obvious errors and fallacies, my argument is more directed towards those who cite Hume as having answered Aquinas, or who try to use Hume's arguments to answer Aquinas. Even by the standards of his age, Hume's arguments are weak; compared against the greater philosophers of earlier times, they are wholly unsound.

12.3.1. The Cosmological Argument

I will start with Hume's 'answer' to the first way. Now, it is not obvious that Hume is referring to the first way: since he discusses causes, is this not a second way argument? We must remember that what Hume means by 'cause' is closer to what Aquinas means by 'mover' than what Aquinas means by 'efficient cause'. Thus this is closest to the first way. In this response, I shall use Hume's terminology (except when citing Aquinas), and thus when I say 'cause' I mean what Aquinas referred to as a 'mover'.

> But if so many difficulties attend the argument a posteriori, said Demea, had we not better adhere to that simple and sublime argument a priori, which, by offering to us infallible demonstration, cuts off at once all doubt and difficulty? ...
>
> You seem to reason, Demea, interposed Cleanthes, as if those advantages and conveniences in the abstract argument were full proofs of its solidity. But it is first proper, in my opinion, to determine what argument of this nature you choose to insist on. ...
>
> The argument, replied Demea, which I would insist on, is the common one. Whatever exists must have a cause or reason of its existence; it being absolutely impossible for any thing to produce itself, or be the cause of its own existence. In mounting up, therefore, from effects to causes, we must either go on in tracing an infinite succession, without any ultimate cause at all; or must at last have recourse to some ultimate cause, that is necessarily existent: now, that the first supposition is absurd, may be thus proved. In the infinite chain or succession of causes and effects, each single effect is determined to exist by the power and efficacy of that cause which immediately preceded; but the whole eternal chain or succession, taken together, is not determined or caused by any thing; and yet it is evident that it requires a cause or reason, as much as any particular object which begins to exist in time. The question is still reasonable, why this particular succession of causes existed from eternity, and not any other succession, or no succession at all. If there be no necessarily existent being, any supposition which can be formed is equally possible; nor is there any more absurdity in Nothing's having existed from eternity, than there is in that succession of causes which constitutes the universe. What was it, then, which determined Something to exist rather than Nothing, and bestowed being on a particular possibility, exclusive of the rest?

> External causes, there are supposed to be none. Chance is a word without a meaning. Was it Nothing? But that can never produce any thing. We must, therefore, have recourse to a necessarily existent Being, who carries the reason of his existence in himself, and who cannot be supposed not to exist, without an express contradiction. There is, consequently, such a Being; that is, there is a Deity[15].

Now, it should immediately be apparent that this is not the cosmological argument, and that it has been parroted by almost every English speaking atheist philosopher since Bertrand Russell as though it were the argument does not make it any less of a straw man. The differences between this argument and Aquinas' are many, and I list just the most obvious. Firstly, the argument is introduced as *a priori*; however the cosmological argument is *a posteriori*, as noted earlier (and is obvious from the start: *It is certain, and evident to our senses, ...*; no *a priori* argument begins with an appeal to observation). Secondly, the first premise of Hume's argument, *Whatever exists has a cause,* is ludicrous. The whole purpose of the cosmological argument is to show that there is at least one thing that exists but does not have a cause. Some modern commentators have even tried to use this contradiction in their arguments against God; but that is absurd. The contradiction just arises from the ill-chosen first premise. Note that the premise of Aquinas' argument is not even *Everything that begins to exist has a cause*, something which may be true but is not immediately self evident. Rather, Aquinas starts with two observations, that *some* things are in motion (or are being changed), and that anything responsible for that change must itself be in a state of actuality. Aquinas argues from the premise that *some things* have a cause to conclude that there must also be *some things* without a cause (Aquinas later shows that there can only be one thing without a cause, but that is not itself part of the cosmological argument). Both of these premises are reasonably apparent, and neither of them contradict the conclusion. I know of no reputable theologian who has argued that *everything* has a cause, and how Hume managed to convert the word *some* into the word *everything* while thinking that he could get away with it has always astounded me; even more astounding is that so many atheist admirers, including some otherwise exceptionally smart people, were fooled by this obvious error.

The argument against an infinite succession differs from Aquinas', and in particular introduces the idea of the totality of all the chain, which Aquinas does not need. This argument is another straw man; if Hume manages to defeat this argument, then that does not prove that its conclusion, that an infinite chain is impossible, is false.

Finally, I wish to highlight that Hume's definition of God, as something whose non-existence implies a contradiction, differs from Aquinas's being of pure actuality, namely something, if that type of being exists, cannot be in a state other than what it is (including cannot come into or out of existence). Hume confuses necessary with *logically necessary*; there is a subtle distinction between the two which the medieval writers were aware of but the less careful Enlightenment philosophers missed.

So having established his straw man, how does Hume argue against it? Firstly, he objects to *a priori* arguments to prove the existence of something. It is questionable whether his analogy to other matter is valid, since God differs from everything else, and thus might well differ in this respect as well. *Whatever we conceive as existent, we can also conceive as non-existent. There is no being, therefore, whose non-existence implies a contradiction.* But our conception is not always logical; we can conceive of things that are in practice contradictory; equally we cannot fully conceive of God due to our own limitations. But, in any case, whatever its merits, this objection fails because the cosmological argument is *a posteriori* and not *a priori.* Equally, this argument responds to Hume's straw man definition of God, and not the classical definition.

---

[15]Hume, *Dialogues Concerning Natural Religion*, IX.

Next, Hume labours this point: *It will still be possible for us, at any time, to conceive the non-existence of what we formerly conceived to exist; nor can the mind ever lie under a necessity of supposing any object to remain always in being; in the same manner as we lie under a necessity of always conceiving twice two to be four. The words, therefore, necessary existence, have no meaning; or, which is the same thing, none that is consistent.* This, it seems to me, is false, and the reason is that Hume does not have a well thought out philosophy of change. As soon as one notes that things may change, and yet remain the same type of being, or change in some ways or not others, then we are forced into an account equivalent to Aristotle's of potentiality and actuality, even if we express it slightly differently, or use different words. Now, if most things are a mixture of potentiality and actuality, and some have more potentia and others fewer, then the idea that something could in principle have no potentia and thus be pure actuality is clear. And such a thing is not capable of change; since ceasing to exist is a type of change, such a thing cannot cease to exist. This is as certain as saying that twice two is four. The only question is whether there is, indeed, a being of pure actuality: and this is what the argument from motion attempts to demonstrate. Again, Hume makes the mistake of arguing about what our minds think. What our minds think is often irrelevant to what is, and this is certainly true if, like Hume, and unlike Aristotle and his disciple Aquinas, we adopt a sloppy formalism which easily allows us to slip into error.

Next, Hume questions, *But further, why may not the material universe be the necessarily existent Being, according to this pretended explication of necessity?* Hume offers another straw man in response to this. One clear answer is that the material universe is not a single being, while every link in the cosmological argument's chain of causes, including its termination or the first cause, is an individual being. More carefully, we can say that the universe is clearly not a state of pure actuality, so cannot be the terminus of the argument from motion. Hume then elaborates a little about his conception of God, and thus demonstrates that he either has not read or has not understood the first part of the *Summa Theologica*, because his idea of the divinity is very different from what classical theists actually believe.

Hume then attempts to refute his own argument against an infinite succession, and fails: *Did I show you the particular causes of each individual in a collection of twenty particles of matter, I should think it very unreasonable, should you afterwards ask me, what was the cause of the whole twenty. This is sufficiently explained in explaining the cause of the parts.* There is a huge difference between twenty and infinity. It is not even possible in principle to show the causes of each individual in a collection of an infinite number of beings. Indeed, if one has a collection of twenty interconnected beings, each with a cause, then there must be between one and twenty beings outside that ensemble. Thus even in Hume's argument he fails to make the case that there can be a set of beings without at least one external cause, given that they cannot all be mutual causes of each other. At least one? Isn't that enough to avoid the conclusion of a single God? However, the cosmological argument itself doesn't claim to demonstrate the unity of God, although (as we shall see) that is a necessary consequence of it which can be established through further reasoning. If one tries to argue in this way, one would be making the fallacy of arguing against a argument that doesn't make the point you reject, while ignoring the argument which does demonstrate it.

Hume continues with an appeal to ignorance (maybe there is something we don't know about), while Aquinas just appeals to what any explanation of physical or immaterial beings must satisfy; and Hume concludes with a statement that generally people haven't found this argument convincing. The fallacy of an appeal to popularity is bad enough; the fallacy of an unsubstantiated appeal to popularity concerning a straw man is even worse.

Thus Hume utterly fails to even address, let alone refute, the argument from motion.

12.3.2. The Argument from Design

So what of the teleological argument? If anything, Hume's discussion of this is (as a refutation of the classical arguments of Aquinas) even worse (although, of course, Hume wasn't responding to Aquinas, but only the writers of his own time; the problem is that Hume has been put to use as though he did criticise Aquinas). At least his presentation of the cosmological argument bore some similarity to the actual argument used. Here is Hume's presentation of the teleological argument, which he then attempts to refute:

> Not to lose any time in circumlocutions, said Cleanthes, addressing himself to Demea, much less in replying to the pious declamations of Philo; I shall briefly explain how I conceive this matter. Look round the world: contemplate the whole and every part of it: You will find it to be nothing but one great machine, subdivided into an infinite number of lesser machines, which again admit of subdivisions to a degree beyond what human senses and faculties can trace and explain. All these various machines, and even their most minute parts, are adjusted to each other with an accuracy which ravishes into admiration all men who have ever contemplated them. The curious adapting of means to ends, throughout all nature, resembles exactly, though it much exceeds, the productions of human contrivance; of human designs, thought, wisdom, and intelligence. Since, therefore, the effects resemble each other, we are led to infer, by all the rules of analogy, that the causes also resemble; and that the Author of Nature is somewhat similar to the mind of man, though possessed of much larger faculties, proportioned to the grandeur of the work which he has executed. By this argument a posteriori, and by this argument alone, do we prove at once the existence of a Deity, and his similarity to human mind and intelligence[16].

To be fair, this argument from design has been used by Christian theologians; most notably Paley, and to this day with the contemporary Intelligent Design movement. I will leave its present day supporters to defend it against Hume's criticisms. However, that it is not the same argument as Aquinas defended is obvious.

Firstly, Hume is reliant on the mechanistic view of nature. This Aquinas denies: the doctor of the Church did not believe and would have contended against the idea that the universe was one great machine. Secondly, the argument from design relies on the harmoniousness of the parts; the teleological argument, on the other hand, is based on the observation that some causes have a limited number of possible effects. Whether the final causes of different beings are well adjusted to each other or are in complete disharmony is irrelevant to Aquinas' argument. For example, Paley famously illustrated his version of the argument by comparing a stone with a watch; to him the stone implied nothing, while the detailed mechanisms of the watch necessitated a watchmaker. Aquinas, on the other hand, can throw away the watch – he has no use for it (except to tell the time); for him even the simplest and most common lump of rock is good enough to demonstrate the existence of God through teleology (and modern physics would have just made it easier for him). For even a stone exemplifies final causality, and that is all Aquinas needs (does a rock lying on a ground exhibit final causality? Yes; if nothing else from the observation that it scatters some wavelengths of light and absorbs others). Thirdly, this argument is based on an analogy with a human made machine, while Aquinas' argument, though it may be illustrated by an analogy (as Aquinas did), is based on purely metaphysical assumptions; namely the existence of final causality, and that directionality is only ultimately possible if there is a guiding will. Fourthly, the argument from design looks at specific examples in the natural world; the teleological argument is based on the existence of final causality itself, a principle

[16]Hume, *Dialogues Concerning Natural Religion*, II.

which, certainly once we have disposed of the mechanistic world view, must underlie any expression of physics. Thus it is not at all clear that a response to the argument from design will be usable against the teleological argument.

Perhaps the difference between the design and teleological arguments are best illustrated by looking at 'God of the Gaps' arguments, which I define to be one which takes the form:

(1) I cannot conceive how this could have happened through any natural process.
(2) Therefore it is inconceivable that this could happen through anything but a supernatural process.
(3) Therefore God.

The fallacy of this argument is obvious; the jump from step one to two (from *I cannot conceive* to *It is inconceivable*) is far too large, and the jump from step two to three may also be problematic. Note that the first jump would not be problematic if it had been proved with certainty that something could not happen naturally. But where this argument is wielded, it is not used in regards to some theory where the science is established beyond doubt, but one where we still have numerous unanswered questions. The God of the gaps argument is an appeal to our ignorance, which then asks that God is used the place of knowledge. Clearly this is problematic, if some scientist later spoils things by coming up with some natural process that explains the harmony. But Aquinas' argument is not an appeal to ignorance, but instead an appeal to what we already know. The universe appears to be rational and well ordered; this, according to Aquinas and his supporters, is enough to demonstrate final causality. Disorganised matter does lack any intelligence, and therefore cannot arrange its own purpose. Indeed, ignorance is the enemy of Aquinas' argument. While we still have ignorance about a certain part of the cosmos, then it may be possible that in that instance the laws of causality break down (at least if we neglect the strong theoretical arguments for universal causality). But as soon as we find an explanation in terms of some natural process, then we bring final causality back in, and add weight to Aquinas' conviction that everything is governed by final causality. The God of the gaps argument breaks down if the gaps are filled; Aquinas' argument breaks down (in a sense) if the gaps are incapable of being filled, and thus the chain of final causality (as witnessed by science) might be broken. The scientist, by pushing Cleanthes out of the door of the house, simultaneously pulls Aquinas in through the window.

The main fallacy of those who wield the argument of the shrinking gaps against God is that they assume that explanations from physical law and explanations through appeal to God are exclusive. The argument runs *We now have a law to explain this, therefore God was not responsible* and so on, until we find that God is not responsible for anything and there is thus no need to invoke Him at all. However, physical law is simply a mathematical description of how physical systems develop in time. This then leads to the question of why that description is successful. Why do material objects obey the laws of physics? For the theist the laws are a description of how God sustains the universe under normal circumstances. Thus, for the theist, to say that something is accurately described by physical law is precisely to invoke God as the author of that event. When the gaps in our knowledge close, it does not force God out of the picture; if the theist is correct it brings God more into the picture. We not only know that God was the actor, but we have a description of how God acts. So the success of science in reducing our ignorance only reduces the involvement of God if we presuppose that the laws of physics operate independently of God. But, to disprove the theist account, that needs to be demonstrated rather than assumed. Saying that the formulation of the laws makes no mention of God is not sufficient; for the theist the laws themselves describe God's involvement and are where God enters the picture. Indeed, Aquinas' fifth way is an argument against the presupposition that the laws operate independently of God, suggesting that final causality (which is to an Aristotelian metaphysics what the laws of physics are to mechanism) has to originate in an immaterial will, namely the will of God; and thus the regularities unveiled by physics are themselves

evidence for God. Thus the shrinking scientific gaps are no argument against God, but precisely what the theists claim when they state that God is sovereign over everything.

Let us look at the most obvious example: biological beings are adapted to their environment in incredible detail and diversity. 'Aha!' Says Cleanthes. 'God!' 'Not so,' replies Darwin, 'that is precisely what we expect from mutation and natural selection. Those alone are enough to guide life into an equilibrium state where everything is well suited to pass on its genes. God is unnecessary.' But then Aquinas steps in, 'The very idea that those genes which allow their organisms to produce viable offspring are *more likely* to be inherited by the next generations demonstrates directionality in nature: some outcomes are preferred over others. This is an example of final causality. Therefore God.'

That final 'Therefore God' in this sketch would, of course, not be such an abrupt jump from the previous statement were Aquinas himself making the argument instead of my brief caricature of him, but instead demonstrated by numerous small but rigorous and immensely detailed logical steps while responding to all the strongest objections (that was Aquinas' style); my sketch is intended as a summary only. My own feelings are that final causality in physics makes a stronger case for the fifth way than in biology (there are, of course, two different senses of final causality in nature equilibrium state teleology, as in this biological example, and direct teleology, as in particle physics; though the biological teleology depends on and thus ultimately implies the physical teleology). Also, Aquinas still requires his second assumption: that only a will can lead to directionality in nature. But, to the small extent that it is relevant for the fifth way, the theory of evolution by natural selection supports Aquinas' argument rather than contradicts it.

So what of Hume's attack on the argument of design? Do any of the responses he offered address the fifth way? I list only a few of the stronger arguments (since it would be too much for me to discuss everything that Hume wrote); other arguments either build on these or are even further from Aquinas' fifth way.

- He questions whether arguments based on analogy are reliable. The fifth way does not rely on analogy.
- He questions about whether experiences of similar events are enough to justify attributing similar causes. The fifth way does not require comparisons between similar events.
- He questions whether we can apply a rule derived from one part of nature to a different part of nature (as the harmoniousness of the universe differs from the harmoniousness of a watch). Did Aristotle derive final causality from one part of his studies, and thus apply it to everything? There might be a historical case for this; I have seen accounts that Aristotle developed his metaphysics from his biological observations, and then applied it to physics. But that the genesis of an idea came from a dubious source is not to say that the idea itself is dubious. Aristotelian causality is a metaphysical proposition. It makes certain predictions about the nature of the physical world, and it can be tested by comparing these predictions to the real world. In particular, it explains why the universe is rational; why there are conservation laws in physics, and so on. Now, of course, we have not explored every crevice of the universe; we cannot say on the basis of experiment that the laws of causality (or more generally the laws of physics) apply everywhere. However, we can calculate some consequences if the laws of physics are in some places irrational or different; and we do not observe these consequences. For this reason, we expect the laws of physics to be universal. If the laws of physics are universal, then so are the laws of causality (which are more fundamental; these may be universal even if physics isn't). Thus I do not think that Hume's point can affect final causality and the fifth way; at least the burden is on the critic of causality to point to some point in the universe where physics does not apply; and if he finds one to show that some other expression of causality does not apply

there as well. What of Aquinas' second premise, that bare matter cannot show a purpose, but only a will? This again is not really based on looking at similarities in nature; it more comes from the definitions. Our experience of our own will only helps us to understand what a will implies, and alerts us to the possibility of a rational intelligence guiding nature. But Aquinas' argument does not in any way depend on our own will; the argument would still stand even if there were no animal life to provide this illustration.

- Hume next questions whether human intelligence (being limited to our planet) is a good guide to what we might expect any intelligence elsewhere in the universe to resemble. The argument from design implies that God is similar to us, and depends on such similarity for its force. But why should we expect non-human intelligence to be so similar to our own? But again, no orthodox Christian would say that God's mind resembles our mind in every detail. We see isolated points of similarity, but that is all. Indeed, orthodox theologians spend more time emphasising how unlike us God's mind is. The fifth way places no such burden on God; He may be similar to us, or vastly different. It does not require that God gives direction in the same way that an archer gives direction to the arrow; merely that there is something which in some way gives direction to nature. Again, Hume's argument is irrelevant to Aquinas'.
- The argument from design does not imply the infinite attributes (omnipotence, omniscience, unity, etc.) usually given to God. However, Aquinas' arguments do imply these attributes (as he showed in subsequent articles of the *Summa*).
- The universe shows imperfect design: 'Thus, the winds are requisite to convey the vapours along the surface of the globe, and to assist men in navigation: but how oft, rising up to tempests and hurricanes, do they become pernicious? Rains are necessary to nourish all the plants and animals of the earth: but how often are they defective? how often excessive? Heat is requisite to all life and vegetation; but is not always found in the due proportion.' However, this is not a problem for Aquinas' argument, which makes no mention of the harmony of nature. Of course, one still has to answer why a God such as the one that Christians worship would make a universe in such a way, but this is a separate issue to whether or not the teleological argument is valid. I will address this at the appropriate point.

Hume presents a strong case against the argument from design, and anyone who supports this argument needs to find an answer to him. But I need not do so here, because I do not defend the argument from design, but instead the teleological argument, which is very different. Hume has nothing to say against the teleological argument, at least not in his *Dialogues Concerning Natural Religion*. He simply does not address it.

Did Hume answer the classical arguments for God? He did not even come close. For the most part, he did not even address them. This, of course, is not Hume's fault; it was not his intention to respond to the particular arguments I raised against him but rather some of the arguments that were being raised in his own day, for example by Descartes and Samuel Clarke. The problem is that modern writers have failed to appreciate that the early modern philosophers Hume was responding to were raising arguments that were very different from those of the classical philosophers. They just assumed that because of a superficial similarity, Hume had dealt with the classical arguments along with the early modern ones. As we have seen, that was by no means the case. Any modern writer who claims that Hume defeated Aquinas has obviously either not understood Aquinas or not understood Hume or both. In this battle which we have forced upon them, Aquinas is most certainly the victor.

12.4. Kant

12.4.1. Definitions

So now let us turn to that other giant of the enlightenment, Immanuel Kant. I will discuss his arguments as they appear in the *Critique of Pure Reason*. Kant's procedure is to discuss and dismiss the ontological argument, and then argue that the cosmological and design arguments are dependent on the ontological argument, and so they must also be false. My main response will be to attack Kant's assertion that the cosmological and design arguments rely on the ontological argument. However, before doing so, I need to discuss Kant's primary reason for rejecting the ontological argument, because if his account is valid then it raises some obstacles which I will need to be careful to avoid.

Before we begin, we need a few of Kant's definitions.

(1) Propositions consist of a subject and a predicate. For instance, in the sentence 'A dog has four legs', the dog is the subject and 'has four legs' is the predicate. The predicate describes properties of the subject.

(2) An idea is a concept that we keep in the brain. Kant argues that it is impossible to completely represent objects by means of a pure concept in the understanding because anything in our thoughts would lack the objects actual existence. Thus there will be a difference between the object in our thoughts and the object as instituted in matter.

(3) An ideal is an individual idea, which is determinable or determined by the idea alone. This seems similar to Plato's concept of forms, though in practice it is rather different. The way this works in Kant's thought is that for every possible predicate, the ideal must either have or not have the property. So we can in principle go through the list of properties, tick each of them off as either applying or not applying as appropriate, and come up with a unique object. This is the ideal. It is completely determined. Of course, it exists in our thought only, and we can only see whether it exists in reality by finding an actual being, measuring all its properties, and seeing if it matches. In practice, this is impossible, since there is an infinite number of predicates, and we can't test them all. Thus, 'The proposition really means that, in order to know a thing completely, we must know everything that is possible, and thereby determine it either affirmatively or negatively. This complete determination is therefore a concept which in concreto can never be represented in its totality, and is founded therefore on an idea which belongs to reason only, reason prescribing to the understanding the rule of its complete application.' The ideal, then, does not describe anything in reality, but things in reality can be similar to it, some closer than others.

(4) Now, we can also think of a ideal which is not derived from anything else, which possesses all positive predicates.

> Thus all the possibility of things (or of the synthesis of the manifold according to their content) is considered as derivative, and the possibility of that only which includes in itself all reality as original. For all negations (which really are the only predicates by which everything else is distinguished from the truly real being) are limitations only of a greater and, in the last instance, of the highest reality, presupposing it, and, according to their content, derived from it. All the manifoldness of things consist only of so many modes of limiting the concept of the highest reality that forms their common substratum, in the same way as all figures are only different modes of limiting endless space. Hence the object of its ideal which exists in reason only is called the original Being (ens originarium), and so far as it has nothing above it, the highest Being (ens summum), and so far as everything as conditioned is subject to it,

> the Being of all beings (ens entium). All this however does not mean the objective relation of any real thing to other things, but of the idea to concepts, and leaves us in perfect ignorance as to the existence of a being of such superlative excellence[17].

I find this passage rather hard to understand, but this is my best attempt to explain it: Many predicates are derived from some other quantity; for example to think of something surviving for a specified time period we first of all need to think of the concept of surviving, which is best explained in terms of continuing for eternity. To survive for a specific number of years is thus a distortion of this original concept. We can only think of to be composed of parts after we have thought of something simple, and then to be be composed of parts means to be other than being simple. Simplicity comes first because each part can ultimately be reduced to a single idea. God (which Kant shortly afterwards identifies with this original being) is thus that ideal whose predicates are not ultimately derived from or a distortion of or a negation of other predicates. Thus we bring together all possible predicates, select those which lack any imperfection, and this is the most real being, or God.

Before proceeding, we need to compare Kant's assessment with the metaphysics of Aquinas. There are a number of important differences. To start with, Kant's ideal is defined in terms of the sum of all its properties. We start with the properties and get the ideal. However, Aquinas works the other way; we start with the form – the abstract concept – and then from that determine its properties. We should not confuse a being's properties with an understanding of what it is – but this is precisely what Kant does.

Kant's ideal is completely determined. The form of Aristotelian logic, however, is not, but describes a large number of possible states. This of course, allows for essential properties, which are shared by all states in the form, and accidental properties, which are shared by only some of them, and are not present in others. Can we identify Kant's ideal with one of these individual states? Certainly not when we introduce quantum physics: properties are described by the eigenstates of operators, but some of these operators do not commute. If an electron has a positive z component of its spin, then its x and y components are undetermined. Yet these are all predicates of the electron. Therefore, we cannot go through the list of all predicates, and say 'yes' or 'no' to all of them. Sometimes, if we say 'Yes' to one of them, we cannot determine others. In this way, Kant's conception of the ideal does not stand in light of modern physics. Can it be partially rescued, by expanding the possible answers to whether a predicate applies to a being to 'Yes', 'No', and 'Undeterminable given our other answers'? This might be a reasonable approach, if it did not undermine Kant's assertion that an ideal can and must be completely determined, i.e that an ideal cannot have accidental properties (and this is a key premise of much of Kant's argumentation). As another example, consider the predicate 'This object is red', and also the predicate 'This object has all its electrons in the ground state'. Now, we know that colour emerges when an electron is excited by an incoming photon, and it subsequently drops back down into the ground state. For something to be red requires a transition between an exited state and lower energy state. Thus, in some sense, we cannot describe something as having a colour while keeping its structure fixed. The energy levels of the electrons must be an accidental property of the substance, and thus is nature cannot be completely determined and fixed.

Note that Kant requires that the ideal is completely determined, because his assessment that it cannot correspond to something in nature and thus just belongs to the mind depends on it. It is quite reasonable to suggest that a single, point like, structure, the ideal, will not find an exact analogue in the real world (unless we have independent reasons to suppose that such a being must exist). However, if the form extends over a region of possible states,

[17]Immanuel Kant, *Critique of Pure Reason*, Part II Chapter III Section III.

which must be the case if it has accidental properties, then it is quite plausible that real-world beings correspond to somewhere within the form; and by thinking about the form we are also thinking about reality. In this way, Kant's idea of a separation between the world of physical phenomena and the world of mental images could be said to rely on his idea that ideals can be completely determined. If that hypothesis falls down (as I believe that it does), then so does almost his entire philosophical system. As with Spinoza, Locke, Descartes and the rest of them (though not Hume, whose ideas concerning causality are incompatible with even classical physics), Kant's beliefs are quite reasonable (though not certain) if the world was governed by Newtonian laws, but since it isn't they are highly questionable, and I do not believe that Kant's metaphysics is of any use to us in our quest for the truth.

However, even Aquinas would, I think, have had difficulties about treating one of these states in isolation, because each state contains the other states of the form as potentia. Kant's system does not allow for change (or motion, in Aristotle's terminology) in an ideal: it is completely static, and has to be, because any change would lead to the change of one of its predicates, and if we do that, we have a different ideal. However, it is clear that beings do change in certain respects, and yet remain the same being. In other words, to think of just one state (or ideal) is not to understand the being in its entirety, and thus we cannot get a complete understanding of a being by just considering a single ideal (in general; obviously this would not apply for a being of pure actuality). To do that, we need the broader concept of form. In this sense, we cannot think of an ideal as completely determined as a useful concept: every being, baring the state of pure actuality, has accidental properties.

Now, beings are determined by form and matter. When we compare two beings, we compare their form; and when we say two beings are the same, what we mean is that they have the same form. Now, to say that two beings have the same form implies that they are identical in all their essential properties, but they might differ in their accidental properties. For example, a material object might be instituted in matter, or as an abstract concept in a mind, or not instituted at all. Thus the mode of its institution is accidental to the form, and (for all material beings) it must be so. It is not like the example of a ball which can be either red or blue; we can narrow down the possible forms by specifying 'a red ball' or 'a blue ball'. But form in itself is completely independent of matter, and this means that which piece of matter in which the form is instituted is accidental to it, and in that case equally whether the form is instituted in matter or the mind. In this way, Kant's argument that we cannot exactly hold the concept of a material being in our mind falls short: we can hold a representation of the same form, with all its essential properties in our mind as is instituted in reality.

Then we come to Kant's understanding of God, and again it is clear that this differs significantly from Aquinas' state of pure actuality. Once again, we see Kant defining God through certain properties that God is alleged to possess. Indeed, when we combine this with Kant's idea of complete determination, we get an ideal of God whose properties are fully known. This is very different from Aquinas' approach: as before, Aquinas starts with a form, understood in only the most elementary sense, and then derives some properties from it. Aquinas does not require a complete understanding of God, indeed Aquinas held this to be impossible. Rather, we can understand something about God partly through analogy and partly through negation; for example, for Aquinas, we understand God's simplicity as the negation of being composed of parts. We understand beings (in part through observation) which are a mixture of potentiality and actuality (which includes being composed of parts), and it is only on account of this we can ponder a little of what it means to be purely actual. Another important difference is that Kant views God in terms of ideals; conceptions that only truly live in the mind (although they might resemble real, extra-mental beings); Aquinas views God as a real being, external to our minds, which we can understand a little about by studying the natural world.

Thus, I think that Kant starts out with a flawed metaphysics and set of definitions. But we should proceed onwards, and see where he goes with this.

12.4.2. The Ontological Argument

> If we admit the existence of something, whatever it may be, we must also admit that something exists by necessity. For the contingent exists only under the condition of something else as its cause, and from this the same conclusion leads us on till we reach a cause which is not contingent, and therefore unconditionally necessary. This is the argument on which reason founds its progress towards an original being.[18]

This is a crude statement of a cosmological argument. Kant then asks which of our ideals is consistent with this idea of existing by necessity. He states that the concept which does not seem to depend on any other concept, the God that he described earlier, best fits this description. Indeed (as Kant develops this argument), it is the only thing which could fit this description. There is some confusion in his argument here between the concepts of causality (which lies behind his original being) and explanation (which lies behind his highest being whose concept is not derived from something else), but let us continue.

> This therefore is the natural course of human reason. It begins by persuading itself of the existence of some necessary Being. In this Being it recognises unconditioned existence. It then seeks for the concept of that which is independent of all condition, and finds it in that which is itself the sufficient condition of all other things, that is, in that which contains all reality. Now as the unlimited all is absolute unity, and implies the concept of a being, one and supreme, reason concludes that the Supreme Being, as the original cause of all things, must exist by absolute necessity.[19]

This is Kant's expression of the ontological argument, and the background for his discussion of the traditional arguments for God.

Kant asserts that there are only three possible ways to prove God from reason alone: firstly, an a priori argument (the ontological argument), secondly an argument based on indefinite experience (i.e. laws of causality, and so on: the cosmological argument), and thirdly an argument based on definite experience (how we observe things to be: the argument from design). Like Hume, he makes no mention of the teleological argument, though perhaps that might be incorporated as a form of his argument on indefinite experience.

So what are the objections that Kant raises against the ontological argument?

(1) It is not in practice clear that the absolutely necessary being from causality has to be the highest ideal. Some more limited being might have the properties of unconditioned necessity. Thus he struggles to see that the qualities of a necessary being (as obtained from causality) can be understood, and since Kant understands beings only inasmuch as he completely understands all their properties, this means that it is difficult to understand what a necessary being is.

(2) The concept of an absolutely necessary being is seen as a concept of pure reason, just an idea, which is only needed to give a completeness to our system of ideals. Just because reason requires it does not prove its objective existence.

(3) We can define a necessary being as one whose non-existence is impossible; however this definition does not help us to understand why it is necessary or its non-existence is inconceivable, or indeed, what such a thing is. Indeed, people argued from necessary truths (such as from geometry) that the concept of a necessary being makes sense, without questioning whether the concept itself is intelligible.

---

[18]Immanuel Kant, *Critique of Pure Reason*, Part II Chapter III Section III.

[19]Immanuel Kant, *Critique of Pure Reason*, Part II Chapter III Section III.

Obviously, there is a difference between a necessary truth of logic and a necessary being, so the analogy is not apt. The idea of the non-existence of a necessary being is not contradictory. For example, we can say that triangle (a subject) has three sides (a predicate). If we deny the predicate, we have a contradiction: the triangle's three sides don't exist. However, if we deny both the subject and the predicate: the non-existent triangle's three sides don't exist, there is no contradiction, since only that which exists can have a existent predicate. In the same way, the statement 'a existent in reality concept of a necessary being does not exist' is a contradiction, but 'a non-existent in reality concept of a necessary being does not exist' is a tautology. Thus there can never be a contradiction if we remove the predicate and the subject together.

(4) If we grant the idea of the most real being possible (i.e. non-contradictory), then it seems that for this being only the removal of its predicate is contradictory. However being not contradictory is only a sufficient but not a necessary condition for something's possibility. However, the concept of a necessary being existing is either just a tautology (in which case the argument gains nothing), or nonsense. Every proposition involving existence is synthetic (i.e. not a matter of reasoned analysis alone), so it is impossible to say that removing the predicate of existence is non-contradictory, since the principle of non-contradiction applies only to analytical propositions.

(5) Existence is not a logical predicate but a real one; in other words it is not something that applies to concepts in the mind, but only to real objects. Being is not a real predicate: as a concept it cannot be added to the concept of the thing. It is merely the admission of a thing. So while the phrase 'God is almighty' contains a subject and a predicate, so adds something to the concept of God, to say that 'God is' adds nothing to the concept of God. Thus existence is not a predicate in the same way that (say) being almighty or anything else is. A hundred real dollars is of the same worth and value a hundred possible dollars, but only the former is accepted as payment when I try to buy my lunch. While the idea of a Supreme being might be useful, we do not incorporate it into existence merely by thinking about it.

What are we to make of this set of arguments? For the most part, they are just irrelevant, because Kant's definition of God is different from Aquinas'; indeed Kant's whole approach is very different from Aquinas', and he proceeds from a very different set of assumptions. For example, Kant's idea of a necessary being is defined as something which can't not exist without a logical contradiction. Aquinas' being of pure actuality, however, is defined as something which is incapable of change (though capable of changing others). From Aquinas' definition, we see that if it exists it is not capable of going out of existence or coming into existence, but no logical contradiction is implied by it not existing at all (that has to be argued for starting from the existence of contingent beings). Kant views predicates as things which add to a concept; Aquinas views predicates as things which are derived from knowledge of a form. Kant's ideal is static and singular, Aquinas' form is varied, and accepts the possibility of accidental properties. Perhaps most importantly in this context, Kant views an idea as the object of our thoughts, rather than, as Aquinas does, that by which we try to understand the object of our thoughts. In other words, according to Kant, when we think about an object we are not thinking of the object in itself, but only an abstraction of it. However, Aquinas (rightly) denies this: when we think of some being, it is the actual being in reality which is the object of our thoughts (or, more precisely, it is the form of the actual being which is instituted by the being's matter is also instituted – if we make no mistakes – in our thoughts).

So what of the five objections I have highlighted above? The first is simply irrelevant: Aquinas does not try to relate a necessary being to some highest reality in our concepts, since he neither admits to a necessary being as Kant understands the term, nor requires

a previously conceived idea of a most real being to which to compare it. The second is equally irrelevant: Aquinas does not argue that the concept of the being of pure actuality is required to explain the concepts of lesser beings, but only that existence of actual lesser beings implies some real being of pure actuality. Aquinas does not need to describe what a necessary being (as Kant uses the term) is, because he does not call on the concept. Equally the fourth objection bears no relation to anything in Aquinas' *Summa*, or my own work.

The final objection requires a little more attention. I have been treating the non-existent state as one of the potentia of a being: something that it can change into. I have also stated that existence is accidental to the form. In other words, I seem to be treating existence as though it were a property of beings; which seems to bring me into contradiction with Kant's final postulate. I noted this objection following definition 4.5.1. Now, it is clear, given what we know about physics, that things can move in and out of existence; so the idea of making a non-existent state as one of the potentia is an obvious extension. Now, we can think about how predicates are treated in the Aristotelian system I am developing. We start with a form, which is a description of the eigenstates of a connected set of Hamiltonians. From the Hamiltonians, and the operators that correspond to the observables, we derive the accidental and essential properties of the being (which are Kant's predicates). So, when we say that 'Being X is red', we could mean either one of two things. The first is basically a tautology, if redness is an essential property of the form. We don't add anything to the concept of X by saying that it is red, because as soon as we understand the form of X we understand that it must be red (although the statement 'X is red' is not useless, because it may not be immediate apparent that the form implies redness without complicated and lengthy calculation; or indeed 'X is red' might be used as a summary of the conclusion of that calculation). The second possibility is that redness is an accidental property of X. In this case, by saying that X is red, we are saying that one of the potentia in which X has the property of redness is actual. If we think of X without any predicates, then any of its potentia could be actual, while by saying X is red, we reduce the number of potentia which could be actual at that time. Now, the statement that 'X is' acts in the same way. It shows that one of the potentia in which X happens to exist instituted in matter is actual, while those potentia where X does not subsist materially are not actual. Similarly, the statement 'X is not' restricts us to just consider those potentia where X is not existent. In this way, the predicates of existence and redness relate to the being in the same way. Thus, in this system, existence is indeed a predicate. Also, when discussing X, we usually refer to the form of X. That form can be reasoned about and discussed by itself, regardless of whether there is any extant object with the form. So saying that X exists does add something to the concept of X; it tells us that that form is somewhere united with matter to form and existent substance. We are not just discussing something in our imagination, but something that is also out there in the real world. Clearly, for most beings, existence is an accidental property. There may, however, be exceptions (in which all the potentia have the property of existence), and one of these is the being of pure actuality.

Therefore whether existence can be regarded as a predicate depends on the underlying metaphysical assumptions. If we adopt Kant's system, where he defines an ideal in terms of its properties rather than letting the form determine the properties; and treats an idea as existent in the mind rather than that by which we understand an object in reality, then his conclusion may be valid in the sense that it follows from the premises. But this does not mean that his conclusion is truly correct, if his premises are wrong. In the superior Aristotelian system, the metaphysical premises lead us to accept existence as a property of beings; and since I am adopting these premises here, I am right to treat it as one.

12.4.3. The Cosmological Argument

Of course, so far we have been discussing the ontological argument, so it is not clear why it should be relevant for the cosmological argument. That Kant's arguments thus far

have proved to be irrelevant for Aquinas is not to dismiss them: they were not intended to address Aquinas, but rather Descartes' form of the ontological argument, and here I think they are reasonably effective (or would be if the underlying metaphysics were sound, but since the same criticism can be levied at Descartes for much the same reasons, this is not too large a disadvantage: Kant has shown that even if we grant Descartes' assumptions, his arguments don't stand up). This is just the prelude of the discussion we are interested in.

Kant defines the cosmological argument in the following way,

(1) If there exists anything, there must exist an absolutely necessary Being also.
(2) Now I, at least, exist.
(3) Therefore there exists an absolutely necessary Being.

This obviously requires some expansion, and Kant then proceeds to do so, in the following way

> The proof then proceeds as follows: The necessary Being can be determined in one way only, that is, by one only of all possible opposite predicates; it must therefore be determined completely by its own concept. Now, there is only one concept of a thing possible, which a priori completely determines it, namely, that of the ens realissimum. It follows, therefore, that the concept of the ens realissimum is the only one by which a necessary Being can be thought, and therefore it is concluded that a highest Being exists by necessity.[20]

To my mind, this isn't much of a clarification. Now, once again, it is clear that Kant is not addressing the first way, nor any of the five ways, nor, indeed, any of the classical formulations of the cosmological argument. For, *given Kant's definition of a necessary being as a logically necessary being*, the first premise is not something that most of the great Aristotelian Christian theologians would have accepted (although I believe that some of their Islamic predecessors toyed with the idea); certainly it is not what they argued in support of. Kant's necessary being is not the same as the being of pure actuality that the first way leads us to (or, indeed, the concept of a necessary being that I use below). Given this, even if Kant succeeds in demolishing the argument he is attacking, it does not follow that he has succeeded in demolishing the far stronger medieval arguments.

But let's take a look at Kant's objections, and see if there is anything which can be adapted to address the first way. Kant admits that the argument seems to add experience to pure reason, and thus seems to avoid the problems of the ontological argument. However, he believes that this addition is only apparent. In practice, the cosmological argument is a matter of reason only, and the argument only rests on the first premise (pure reason) and does not in fact require or use the second (experience). His arguments for this are:

(1) He claims that the properties of a necessary being cannot be derived from the concept. Thus we need the ontological argument to clarify what the first premise entails. However, as Aquinas showed, many of the properties of a being of pure actuality (and enough to establish it as God) do follow from the basic concept, and from what we know about it given that it is the fulfilment of the five ways. Kant's objection here is thus either a straw man, or false, or both.
(2) He claims that it is assumed that the concept of a being of the highest reality is perfectly adequate to the concept of absolute necessity in existence. This is a straw man, since Aquinas neither requires nor uses these concepts.
(3) He then discusses the problems of distinguishing different necessary beings. Again, this is either a straw man or false, since Aquinas went on to show that there could only be one being of pure actuality[21].

---

[20]Immanuel Kant, *Critique of Pure Reason*, Part II Chapter III Section V.
[21]Aquinas, *Summa Theologica*, I.11.3.

(4) Kant asserts that our understanding of causality is taken only from the observable material world, and that it therefore need not apply when we go beyond the natural into the immaterial. However, this is false. Our understanding of causality is taken from metaphysics. This makes certain predictions about how causality ought to appear in the material world, and we can test each metaphysical system by comparing those predictions against nature. This, for example, allows us to exclude the metaphysics of Kant (who cannot cope with observables related to non-commuting operators) and Hume (who contradicts the conservation of four momenta, among other things), while that of Aquinas survives. It does not by itself mean that Aquinas' metaphysics is right; but it does mean that it is plausible. Now, it is reasonable to suppose that material causality is different between material and immaterial beings, since the presence or absence of matter is by definition the sole distinction between the material and immaterial. However, final and efficient causality are derived from the concept of existence itself, while formal causality, as the distinguishing principle, applies to immaterial beings as much as the material. Causality is a statement of the premise that the universe, and everything within it and outside it, is explicable rationally, either in terms of other beings or (in the case of God) it needs no causes because it is not dependent on anything else. If we demand the rationality of nature and supernature, then that implies at least a weak form of causality. There is nothing in that statement which depends on the beings under consideration being material. Thus causality transcends the gap between the material and the material. If we find, as theists claim, that the material world alone is insufficient to explain all material effects, then, as long as we are to maintain the premise of a rational cosmos, we are forced to consider immaterial causes. Thus there is every reason to believe that these causes apply beyond our material world (at least for any beings related directly or indirectly to our visible world), and this is all that Aquinas (and I) require. Kant thus makes the mistake of assuming that the principle of causality is derived by induction. It is not. It is one of numerous, but not infinite in number, self-consistent metaphysical possibilities, but the only one of these (as far as I am aware) which survives the scientific test of deduction and comparison with nature.

(5) He argues that the arguments against an infinite series again are derived from induction from the sensible world[22], and thus is equally questionable when extrapolated to the insensible. But again, this is false, at least for Aquinas' argument, which is based only on the logic of an *essentially ordered* series. We identify that it is essentially ordered rather than an *accidentally ordered* from the observable world, but logic tells us that an essentially ordered series cannot switch to becoming accidentally ordered when moving from the material to the immaterial, since the two types of ordering are entirely different things. Thus, again, the principle does not arise from induction, but from formulating self-consistent possibilities, and testing them through deduction and comparison.

(6) Kant argues that the concept of the necessary being is incoherent, since by removing every type of condition, we make any concept impossible. Again, this is irrelevant, because neither the concept of the necessary being nor the removal of dependencies on the concept is used in the traditional cosmological arguments.

(7) The principles which allow us to identify the logical possibility of a concept of all united reality (without any internal contradiction) are also based on principles applied to our senses. Again, this is irrelevant, since this concept is not used by Aquinas or any of those of his tradition.

[22] i.e. the world which is open to our senses.

(8) The idea of a necessary being is in itself something that can only be derived *a priori*. Therefore, the cosmological argument is ultimately a priori, and this ontological. This argument, of course, fails for familiar reasons.

(9) The idea of God is wholly incomprehensible. That it is not fully comprehensible to us is admitted by all; but that it is wholly incomprehensible is clearly false in that many people (not least our medieval hero) do or did understand a little about it. And that little understanding is enough to draw some very significant conclusions.

(10) The concepts of necessity and contingency do not apply to things in themselves, but only logical concepts. Once again, Kant rests this on his own definitions of necessity and contingency, which are not shared by the medieval theologians (or myself).

(11) Form was considered by the ancients as contingent, but pure matter (or prime matter) as something necessary. However, 'matter and everything in general that belongs to the world are not fit for the idea of a necessary original Being, as a mere principle of the greatest empirical unity, but that we must place it outside the world. In that case there is no reason why we should not simply derive the phenomena of the world and their existence from other phenomena, as if there were no necessary Being at all, while at the same time we might always strive towards the completeness of that derivation, just as if such a Being, as the highest cause, were presupposed.' Of course, in the scholastic system prime matter is pure potentiality with no actuality, God is pure actuality with no potentiality. They are complete opposites of each other. The rest of Kant's statement here seems to be mere assertion without demonstration; for the cosmological argument purports to show that there is a reason why we should not derive the phenomena of the universe from just the universe, and Kant has not answered this argument[23].

In summary, Kant completely fails to address the actual cosmological argument, but only his straw man.

I need not go on to consider Kant's discussion of the argument from design, only to say that it is just as bad as his discussion of the cosmological argument. It is irrelevant, anyway, since neither Aquinas nor I use the argument from design.

12.5. Summary

None of this discussion should, of course, be seen as a criticism of Hume or Kant. They did not reply to Aquinas' first or fifth ways, but they were not intending to. They were replying to the arguments of their contemporaries, who had tried to twist the traditional arguments to enlightenment principles. Perhaps they would have felt, if they were aware of Aquinas, that his metaphysics had been well refuted already, and the five ways with it. In the light of Newton's successes, that is not completely unreasonable (though not as certain as they believed). In the light of Newton's failures, as we have seen, this supposition seems premature and foolish. However, for a modern writer to cite Hume and Kant or any of their generation as having proved the traditional arguments for God false is a case of either inexcusable ignorance or inexcusable stupidity. They did not refute the five ways; they did not even address them, or any argument similar enough to them for their objections to be considered seriously.

Of course, that is not to say that arguments cannot be made against the five ways. Most critics that I have seen have tried to understand the traditional arguments within their modern world-view, and have therefore failed in the same way that Hume and Kant did. A few, however, have tried to answer Aquinas on his own terms. The best attempt, in my view, of trying to offer a fair critique of Aquinas' work is Antony Kenny (Kenny, 1969), although I feel that Kenny failed, by bringing in too many modern philosophical assumptions into his reading of the Medieval philosopher and misunderstanding Aquinas' argument in

[23]Immanuel Kant, *Critique of Pure Reason*, Part II Chapter III Section V.

too many places. He made too much of the Aquinas' physical illustrations (which were built on a false Aristotelian physics), and spent most of his time refuting them rather than the actual metaphysical arguments (which are more solid). He also objected to Aquinas' dependence on 'outdated' philosophy, but should, as I believe, the Aristotelian concepts be found to be consistent with modern physics, then that dependence is to Aquinas' advantage rather than disadvantage. Responses to Kenny's work can be found in, for example, (Feser, 2009, 2014; Oderberg, 2010; Velecky, 1994; Wippel, 2000).

I should emphasise, that although I use the words 'necessary' and 'contingent' later, my definitions will be very different from Kant's. I shall not subsequently use Kant's terminology, which I don't believe to be useful.

CHAPTER 13

A Classical Theory of God

13.1. Definition of God

DEFINITION 13.1.1. **Existence**:
There are six senses in which something can be said to exist.

- The first, which is the sense used when discussing the existence of God or of some other particular being, is that an existent being is able to interact with other beings; that is if its state is able to be changed by something else (for example by actualising a potentia), or if it can change the state of something else. This is the primary meaning of the word, and unless otherwise stated or obvious from the context, when I discuss 'existence' I will use it in this sense.
- Secondly, we can say that something exists potentially, i.e. as a different potentia of a given form which has a different potentia actual (and existent in the first sense). For example, if an electron in a ground state exists (in the first sense), then the same electron in an excited state can be said to exist potentially within that electrons form.
- The third sense is similar, and refers to virtual existence, i.e. as a possible final cause or constituent of an existent (in the first sense) object. For example, if there is an existent (in the first place) pair of water molecules, then the constituent single oxygen and two hydrogen molecules exist virtually within the water molecule.
- The fourth sense applies to relations between beings; we can say that there is an existent link between two beings (for example, in human relations, fatherhood or friendship can be said to exist in this sense).
- The fifth applies to purely abstract quantities, and here to say that something exists means that it is logically possible given the various definitions and axioms already stated (in this sense, we talk about mathematical quantities existing).
- The sixth applies to abstract conceptions held in the mind, for example imagined images or conceptions, or forms which we grasp with the intellect.

These six meanings of the same word ought to be clearly distinguished. In the sections below, I will use the first of these meanings when I refer to existence. Δ

DEFINITION 13.1.2. **Being**:
A being is an individual existent substance. This might be simple (not made of parts) or composite. So for example, an electron, an atom, an oak tree, a dog and a man are all beings. A collection of five isolated electrons is not a being but five beings; but a proton and electron interacting with each other but otherwise isolated are both beings in themselves and together make another type of being, the composite hydrogen atom. A man's eye or leg is not a being because it is a component part of a substance, the man, but not a substance in itself. A unicorn is not a being because it does not exist; but a particular horse is a being because it does exist.

How can we justify saying that five isolated electrons are five beings, but an electron and proton bound together into a hydrogen atom is a single being? I think that the operative word there is *bound*. Suppose, as seems likely, that electrons are elementary particles; while protons are constructed from three elementary quarks. There is a difference in the nature between an electron and quark on the one hand and the proton and hydrogen atom on the other: the proton and hydrogen atom are composite, while the electron and quark are elementary. There is thus a difference in the nature of the being; one is a composite being, one is elementary. For an elementary being, there is no difficulty in saying that it is an individual substance. A composite being, on the other hand, requires further consideration. We may call it a single being if, in normal circumstances, it requires work to separate it into its component parts (or, alternatively, if there is a stable effective Hamiltonian constructed from creation operators for the various energy levels of that substance). That is, if the potential energy of the system is a local minimum when the elementary matter from which it is constructed are arranged close together in a particular pattern (related to the form of the substance), and larger if they are arranged into some different form or moved so far apart that they no longer interact with each other. In this sense, a wooden table, which is (mostly) strings of cellulose bound together is as much a being as the individual quarks and electrons that the cellulose is made from.

The one slightly uncertain element of that definition is the use of 'normal circumstances'. By this, I mean that in the substance's natural condition, that is the minimum state where it is able to endure over time. For example, two women chained together are not a composite being despite that it requires work to separate them, because being chained together is not the natural condition of women. A lump of sugar in a bowl requires work to break it apart, so it is a single being, even though if we were to place it into a cup of tea it would rapidly dissolve and cease to exist as a lump of sugar but individual sugar molecules. The natural condition of a sugar lump excludes those occasions when it is surrounded by water. Of course, some degree of common sense is necessary to extend this definition to all circumstances. Δ

DEFINITION 13.1.3. **Derived existence**:
A being whose existence is derived originates from another being or set of beings; it has an efficient cause. This derived existence may emerge in several ways; for example it may no longer require its forebearer for its continual existence, as in an electron emitted from an atom as a result of beta decay; or it may only continue to exist as long as that from which it originated exists; for example the music played by a flautist would cease if the flautist suddenly died during the performance. The first of these refers to instantaneous creation; the second a continuous creation; but in both cases the being is created. In short, a being with a derived existence can only exist if something other than it exists (or has existed). Δ

DEFINITION 13.1.4. **Self-existent**:
A being which is self-existent does not have an efficient cause. It does not therefore come into existence. The reason that it exists is either because logically it must exist given its nature and existence (this is not a tautology; there is a difference between *must exist given its nature*, which means that *if* it exists it cannot start or stop existing, and *exist*, which addresses that initial 'if'), or it is in some way contained within itself. Of course, while it is obvious from observation that there are beings with a derived existence, it is not so apparent that there is anything that is self-existent: this needs to be demonstrated. Equally, it is not obvious that there is nothing that is self-existent: this too, if true, should be

demonstrated rather than assumed. A self-existent being does not rely on something other than it to exist. Self-existence and derived existence are mutually exclusive and complete. △

Discussion:
There is one important clarification which needs to be made here. Suppose that A and B are in some respect distinct, and B can only exist if A exists. There are three ways in which this can be achieved. Firstly, B's existence may be logically necessary given A; that is if A exists then B must also exist, and also if B exists then A must also exist. Secondly, we might have the case where if A exists, then B must exist, but the existence of B does not imply the existence of A; for example, there might be several different causes which could lead to B. Thirdly, it may be that, given A, B may or may not exist: if A is a mind, then it is free to choose whether B should exist; if it is a material body there is a certain non-zero probability that it will cause B to exist (perhaps via some process of spontaneous emission) and a certain non-zero probability that B will not exist. The music is only produced if the flautist freely chooses to play. While discussing beings with derived existence, I only refer to the third possibility, where B need not exist given A. Here B is a different being from A and has a different substance. In the first case, where both A implies B and B implies A, A and B cannot be logically separated despite being in some respect distinct; they are therefore of one substance and not different beings but rather distinct aspects of the same being. The second case is between the two; since B is not self-existent, it must derive its existence from something other than B, so this will reduce to either being part of the same substance as A or a derived and separate substance depending on the relationship between a particular example of B and its individual efficient cause. □

DEFINITION 13.1.5. **Contingent**:
A contingent being is one that either may or may not exist; it has the capacity to come into existence and to cease existing. In other words, its existence either depends or did depend on something external to itself. For example, a dog comes into existence at its conception, and ceases to exist at its death. There are therefore times before its conception and after its death when it happens to not exist. It is not necessary, as part of the dog's essence or form that it does exist (the form of the dog remains regardless of whether or not the dog exists; for example when we think of a pet dog who died in our childhood we bring to mind the form of the dog which remains in our memory even though the substance of the dog, the union of form and matter, no longer exists; similarly we may predict the future by considering the form of an artefact that we are yet to build. Form does not cease to exist; but substance can cease to exist if the being is contingent). Even if it had continued eternally, a dog would still be a contingent being because it is part of a dog's nature that it can die and the dog consequently ceases to exist. Everything we see around us is contingent. △

DEFINITION 13.1.6. **Necessary**:
A necessary being is a being that is not contingent. If it exists then it must exist: it cannot come into our out of existence, or, more broadly, it cannot have any boundaries in space or time or any other sense linked to space and time (that we cannot easily imagine boundaries other than in space and time does not mean that there are no such boundaries: we have to include the possibility in our discussion, even if it later turns out to be unrealised), where on one side of the boundary it exists and on the other side it does not. It is possible that the form of such a being can be imagined but have nowhere and at no time existed in practice; but if such a being does exist, then it must always and everywhere exist. A necessary

being cannot have an instantaneous efficient cause (as this would imply that it comes into existence at a particular time), but it can have a continuous efficient cause, for example the tune played by a (hypothetical) eternal violinist. In this way, necessity is distinguished from self-existence. Of course, it is not self-evident that there are any necessary beings: this needs to be demonstrated.

I ought to clarify what I mean by 'boundary'. A 'boundary' is something which divides two natures or two beings; an incompatibility, for two natures at the same level of existence.

A necessary being is defined as one which cannot in principle have an instantaneous efficient cause, which means that it cannot have a beginning in time. Through Lorentz symmetry (which mixes the time and space coordinates), it is also evident that it cannot have a beginning in space. The same argument will apply for other physical properties such as mass and spin, which are linked by symmetries to space and time. What about other types of boundary?

The concept of a boundary may be applied ontologically as well as in space and time. For example, the form of an up quark is distinct from the form of a down quark (I use form as an example because it is not material; this boundary is not a matter in space and time): they are fluctuations in distinct fields which have different masses, electric charges and so on. There is thus an ontological boundary between these two forms: to say that something is both an up and down quark is contradictory. Now let us suppose that there is a unified field which describes both up and down quarks (and for this illustration nothing else); perhaps the quark flavour structure emerges through some symmetry breaking. To be an up quark and to be a down quark are both consistent with being fluctuations in this field. This does not mean that being up quarks and down quarks are the same, or that this field is either an up quark or a down quark or both. It means that the field is ontologically prior to up and down quarks; a third type of being that lacks the boundary between up and down quarks. The difference between the up and down quark is a difference in kind, and this I classify as a boundary. The difference between the more fundamental field and the quark is a matter of relation alone; and this I do not include in the definition of boundary. The quark field is subsistent on the fundamental field, dependent on it and neither separate, identical nor in contradiction to it.

Can a necessary being have the same sort of boundary that distinguishes between two different types of form? As this point in the argument, that has to remain a possibility. I have defined necessity as lacking the possibility of an instantaneous efficient cause; distinction between forms is not a matter of efficient causality. That is not to say that a necessary being must have these sorts of boundary, only that it is not contained within the definition.

Note that I am not defining a necessary being as a logically necessary being (something that exists in all possible worlds or realities, or is inconceivable that it could not exist, or whose non-existence entails a logical contradiction), but as something that if it exists in this reality must exist everywhere and at all times. It may be possible to imagine a logically self-consistent reality where such a being does not exist even though it does in our reality, or self-consistent realities where one or more such beings do exist even though they do not here. A necessary being may also be logically necessary, but it doesn't need to be. △

DEFINITION 13.1.7. **Logically necessary**:
Something is logically necessary if denying it involves a contradiction. Anything that is logically necessary will be necessary in the sense that it cannot come into or out of existence; it is not true that every necessary being is also logically necessary. △

Objection:
The objection starts with the observation that there are terms which are co-referential. For example *David Cameron* and *The Prime Minister of the United Kingdom who succeeded Gordon Brown* are entirely equivalent, and we may (it is supposed) freely exchange them in any sentence. So, for example, *David Cameron has been on holiday with his family in the Caribbean* is entirely equivalent to the statement *The British Prime Minister who succeeded Gordon Brown has been on holiday with his family in the Caribbean.* The ability to extend the language in this way is seen as a important way of allowing us to infer relations between different sentences or concepts; the system of language which allows this replacement is seen as extensional.

The statement *The tallest man in Indiana is taller than anyone else in Indiana* is obviously true. Now suppose that the tallest man in Indiana happens to be one Sam Small. In that case, the phrases 'the tallest man in Indiana' and 'Sam Small' are equivalent and may be used interchangeably. Thus our initial statement is equivalent to saying *Sam Small is taller than anyone else in Indiana*, which is also true.

Now suppose that we repeat the statement with the expression *The tallest man in Indiana is necessarily taller than anyone else in Indiana.* This agrees with the definition of logical necessity given above. Given that *The tallest man in Indiana is taller than anyone else in Indiana* is a tautology, its truth must be valid throughout time, space, or any other boundaries. However, if we now replace 'The tallest man' with 'Sam Small' we find a statement which is false: *Sam Small is necessarily taller than anyone else in Indiana,* which is not correct given that Sam Small will no longer be the taller than anyone else in Indiana once his cousin, Tom Tall, moves to Indiana from Wyoming. In one sentence we regard the language as extensional; in another place, with the same terms and concepts, it is not extensional. This indicates that there is a problem somewhere in our reasoning, and given that a concept cannot simultaneously be extensional (different words can describe the same thing) and not extensional (those same different words cannot describe the same thing) many have suggested that the problem must lie in the concept of necessity. •

Response:
It is surprising to me that this objection has been raised by many well-thought of philosophers. It is essentially a confusion between the concepts of equality and identity; equality meaning (in this sense of the word) that the two objects resemble each other in certain respects; identity that they resemble each other in every respect. The flaw in the argument is either in the concept or use of the co-referential; either the concept is flawed, or it is used in ways beyond those situations where it ought to be used. The statement *Sam Small is taller than anyone else in Indiana* is not equivalent to the statement *The tallest man in Indiana is taller than anyone else in Indiana* given a co-referential. Rather, we should spell it out in full detail:

Sam Small is taller than anyone else in Indiana.

is equivalent to the two statements

The tallest man in Indiana is taller than anyone else in Indiana.

and

Sam Small is the tallest man in Indiana. (1)

Similarly, the false statement

Sam Small is necessarily taller than anyone else in Indiana.

is equivalent to the two statements, the true

The tallest man in Indiana is necessarily taller than anyone else in Indiana.

and the false

Sam Small is necessarily the tallest man in Indiana. (2)

Those who raise this objection seem to be treating the statements (1) and (2) as though they are equivalent and interchangeable, which they are clearly not. The problem is not with the concept of necessity, but with the concept of extensionality as it is used: rather than saying that *Sam Small* and *The tallest man in Indiana* are co-referential, we should say that they are co-referential only in certain well defined circumstances; that the expressions may be used interchangeably in one circumstance does not mean that they can be used interchangeably in every circumstance; so to use the extensionality of language we must also specify the situations where the co-reference is valid. So, for example, we may legitimately say that *Sam Small* and *The tallest man in Indiana at the present* are co-referential, and rephrase the statements accordingly. Then the final contradiction, that *Sam Small* and *The tallest man in Indiana* are both co-referential and not co-referential is resolved without having to question the concept of necessity. The claim that *Sam Small* is interchangeable with *The tallest man in Indiana at the present*, required in the first instance, do not contradict the claim that *Sam Small* is not interchangeable with the phrase *The tallest man in Indiana across all time* required in the second instance. So, we may write,

Sam Small is presently taller than anyone else in Indiana

is equivalent to the two statements

The tallest man in Indiana is taller than anyone else in Indiana

and

Sam Small is presently the tallest man in Indiana. (3)

(1) is the same as (3) if we take the word 'is' in (1) and its accompanying statements to mean 'is presently' rather than 'the concept of Sam Small' is interchangeable with 'the concept of the tallest man in Indiana'. However, we cannot then convert from the statements surrounding (1) to the statements accompanying (2) since 'is presently necessarily', as the phrase would appear in (2), is a contradiction. It is better to have a single meaning for the word 'is' to avoid confusion (A is B means that A is interchangeable with B), and include the modifiers 'presently' or 'necessarily' as appropriate. The flaw that led to the contradiction in the objection then originates from the improper use of a co-referential.

When we use a co-referential, we state it in a way that includes the conditions under which it is valid, and this will enable us to avoid confusion as above. We are better off expressing these co-referential expressions as a syllogism, with the obvious rule that the conclusion is only a necessary truth if both premises represent necessary truths. If either one of the premises is only a contingent truth, then the conclusion would only be contingent. The issue in this example (and those like it) is not a problem with the concept of necessity, but a straight forward logical fallacy hidden by the presentation of writing the argument out sloppily without giving the logical structure in full.

Nonetheless, it was on account of objections such as this that the idea of possible worlds was born, as an attempt to rescue the concepts of necessity and contingency. The idea is to define 'necessary' as being 'true in every possible world'. The response to the objection is similar to what I have proposed above; 'Sam Small' is only interchangeable with 'The tallest man in Indiana' if Sam Small is the tallest man in Indiana in every possible world. Given that he is not, then we cannot use the linguistic trick which begins the objection. However, we do not need to invoke the concept of possible worlds to answer this or other objections to the concept of necessity; merely fully express the objection as a full syllogism, state clearly the conditions under which the statements apply, and watch the apparent contradictions fall apart. 'True in every possible world' is similar to my concept of 'logically necessary', and the differences between the two are not sufficiently large that I need pay attention to them for my purposes here. What I mean by necessity is weaker, and has no analogue requiring a 'possible worlds' model. I do not demonstrate below that God exists in every logically consistent possible world (although that may well be the case); merely that

God must exist in this world given some additional empirical information (such as the laws of causality and that there are some existent objects) which we can know about this world. ○

Discussion:
It is also useful to compare the concept of contingency as defined in the multiple worlds scenario and the notation developed here. In the multiple worlds scenario, if something is contingent, then it appears in some worlds but not all. In particular, if a property b of a substance x is contingent then it appears in some worlds where x (or something corresponding to x) exists, but not every world where x exists. This corresponds (in the definitions used here) to the concept of accidental properties of the form. We earlier defined the form of an object x as the set of Hamiltonians which could describe that object. Naively, we can say that the probability that x has the property b is the proportion of Hamiltonians within the form of x which lead to that property. More accurately, given that the Hamiltonian is of a quantum mechanical system and thus describes an indeterminate system itself, we should write that $P_x(b)$ the probability that a substance built on the union of the form x with matter will possess the property b is given by,

$$P_x(b) = \sum_{H \in x} \omega_x(H) P_H(b),$$

where $P_H(b)$ is the probability that the property b will be realized if the actual object is described by the Hamiltonian H; $\omega_x(H)$ is some weight function describing the probability that the Hamiltonian H will be actual (this may be measured empirically; but, to avoid empiricism, it would be better to estimate this theoretically and then compare against the empirical result), and the sum is over all Hamiltonians contributing to x. Then if $P_x(b) = 0$, we can say that the property is necessarily absent in x; if $P_x(b) = 1$ that it is essential to x; and for $0 < P_x(b) < 1$ it is a possible, or accidental, property of x. These accidental properties are equivalent to a contingent property in a (realist) multiple worlds scenario; although again we have expressed it in language that avoids the need to invoke multiple worlds.

The definitions of contingency and necessity as logically necessary and logically contingent, as implied by this objection and discussion are very common in modern philosophical literature. However, they are quite different from the definitions of the terms which I adopt here (following the usage of the classical philosophers). When an atheist philosopher discusses God as a necessary being, they usually interpret that to mean logically necessary, but that is a much stronger condition than the one used in the classical definition of God. Proving that God is not logically necessary (for example by arguing against the coherence of the concept of logical necessity) does nothing to show that God does not exist. □

CONJECTURE 13.1.8. *It is impossible for there to be something which may cease to exist but did not come to exist. If something can be destroyed, then the process can in principle be reversed and it can be created.*

Proof:
Let us first consider the case of individual elementary particles. It is easy for us to consider the process of destruction. We imagine an object, which has a non-zero probability of decay into two different particles. The non-zero probability means that given infinite time, then it will decay, no matter how small the probability is as long as it is not zero. In other words, an elementary particle cannot last for ever unless it cannot decay into something else. But quantum physics tells us two things: firstly that the probability of three particles interacting with each other is determined by the coupling in the Lagrangian, and this is the same whether one of the particles decays into the other two, or the particle is created

from the other two. Secondly, the laws of physics governing elementary particles are very nearly time reversible (and the *very nearly* is not enough to wreck this argument; it is just a very small fine tuning of the probabilities, not an elimination of the probability). We can therefore switch time into reverse, and say that since there is a non-zero probability of the particle being created, then eventually, given enough backwards time, the particle will at some point be created. As the two probabilities of decay and creation are related, if a particle can decay it can also be created.

Now let us move to complex systems of particles. Here we do have the possibility of equilibrium states and the arrow of time. If the system is in a true equilibrium, then it will always remain there. It either would have continued indefinitely in the past, or come in to being, but this need not concern us, as we are concerned with those states which can cease to exist.

In practice, most objects would be in meta-stable states; that is they remain the same for a time, but at some point decay. In short, there is a small but non-zero probability that it may decay and cease to be at any time (this probability may, of course, depend on the time since its creation; it may even be zero at certain times after its creation, but it cannot be zero at all times after creation, as that is the very definition of meta-stability; the probability will, almost certainly also depend on the object's environment which may change from time to time). These are the types of objects we are interested in. But here the same argument applies as with the elementary particles: if the state is unstable with respect to time going forwards, then it must be unstable with respect to time going backwards; and given an infinite time it would eventually come into being. A similar argument can be applied for those systems which do not even have meta-stable states.

So 'something that ceased to exist but did not come to exist' would be a system which was in a true equilibrium, but at some point moves out of a true equilibrium. But this contradicts the meaning of equilibrium: that state which one can move into but cannot escape, a fixed point of the (universal) Hamiltonian. This is a contradiction.

There is one remaining option: an object whose status depends on the outside environment. In certain conditions it is in a true equilibrium, in the sense that it will endure as long as those conditions remain, but in other conditions it becomes meta-stable or unstable in that there is a non-zero probability that it will decay. Stretching back to infinite time, the external environment was such that it was in equilibrium, but then at some point in time the external environment changes, and it is able to decay. But here, we have to ask about the environment. This must have been in one state (or one small subset of states) for infinite time in the past, but suddenly changed to something else. So was the environment in an equilibrium state (or states if it was doomed to oscillate between one and another which both led to the object being in equilibrium)? Then it cannot have decayed, and the object remains in equilibrium infinitely into the past and infinitely into the future. Was it in a meta-stable state? Then it could not have been so infinitely into the past, and the object could not have been in equilibrium infinitely into the past. In either case, this scenario is also impossible.

We are thus left with the conclusion that everything which can cease to be must also come to be. It is possible for something to neither come to be or cease to be, to remain in equilibrium for all time (the necessary being defined above), or for something to come to be and yet endure forever into the future (so a system in a non-equilibrium state evolves towards an equilibrium state), or for something to never have existed at all; but it cannot last infinitely in past but not infinitely in the future. The same argument applies if we define 'past' as previously in the sequence of causes rather than previously in time; thus we can extent it to non-temporal beings and simultaneous (or even future in time if any such things occur) causes.

We conclude from this that no contingent being, or contingent system of beings can have existed indefinitely into the past, although they may exist indefinitely into the future.

■

CONJECTURE 13.1.9. *There is at least one being which exists necessarily*

Proof:

(1) Everything that exists exists contingently or necessarily.
(2) Everything that exists contingently derives its existence from something else that exists.
(3) This chain is either infinite in the past, circular or it terminates.
(4) If it terminates, the first element in the chain either exists contingently or necessarily.
(5) The chain cannot terminate with something that exists contingently.
(6) The chain cannot be circular.
(7) The chain cannot be infinite.
(8) The chain must terminate with something that exists necessarily.
(9) There are objects which exist contingently.
(10) There is therefore at least one being which exists necessarily.

■

Discussion:
Point 1 is obvious by definition: the definitions of contingent being and necessary being are mutually exclusive and complete. Point 2 is the principle of causality discussed earlier, where it was demonstrated that it holds for every body subject to the laws of physics due to the principle of the conservation of momentum. It is also a more general metaphysical principle, and follows from the assumption of the rationality of nature: if things keep popping in and out of existence without an explanation, then the universe cannot be rational – this irrationality would have spread to the physical world, and it would have been observed. They must be explained, and if contingent, that explanation must reside in something that existed prior to themselves in the chain of causality. Point 3 is again self-evident, although it should be noted that 'past' here strictly refers to the past in the chain of causality, which may be but is not necessarily a temporal sequence. A temporal sequence is one where $A \to B$ implies that A is prior in time to B (or, more strictly, exists in the backwards light-cone); and this is what we usually think of in terms of causality. However, the logic demands we must also consider situations where A follows B in time, or A and B are simultaneous or A and B are non-temporal, i.e. *outside time*, even of the first of these seems to be ruled out by physics and the third is perhaps hard to imagine. However, being hard to imagine does not mean that we cannot think about such things abstractly. 'Circular' means a situation such as $A \to B \to C \to D \to A$, where the chain of causality loops back on itself. This option seems strange, because we usually think of causality in terms of time which is always linear and heading in the same direction; but nonetheless it logically remains an option that needs to be eliminated. Point 4 is basically a rephrasing of point 1, and follows from the definitions.

So far, we have drawn together the premises of the argument; these either follow from definition, logical necessity, or in the case of the principle of causality drawn partly from physics and partly from our assumption that reality is in principle (if not necessarily in practice) understandable by being rational. We now come to the centre of the argument. Point 5 can be argued as follows: if the chain of causality terminates with something contingent, then that object must either come into being from nothing, which violates the premise 2, or

itself extend backwards indefinitely, which is impossible for contingent beings as discussed above (conjecture 13.1.8).

Point 6 follows on the grounds that something cannot be the cause of itself, whether directly (which is clear, since only something actual can be a cause, from the definitions of actuality and existence, but before being caused objects can only be exist potentially), or indirectly through intermediaries. For example, suppose that A caused B which caused C which caused D which then went back in time and caused A. There are two possibilities: either this circle of causes does not interact with anything else, in which case it may as well not be there and we need not consider it further; or at least one of A, B, C and D is the efficient cause of something else (say E) and through that the rest of the universe. But this is the same as saying that the circle of causes A, B, C and D as a whole popped into existence out of nothing, which (if they are physical objects) violates the conservation of four-momentum, and (regardless of whether or not they are physical) violates the assumption that the universe is rational, for in this case there would be no explanation for the sequence of objects $A \to B \to C \to D \to A$; alternatively A could have as its efficient causes both D and something else F that existed in the external universe; but here the circular causes do not help us escape from the chain of causality. Thus, whichever option we choose, the idea of a circular cause either is inconsistent with our axioms, or does not escape the conclusion of this conjecture. I exclude in the discussion the possibility that $A \to B \to C \to D \to A'$, where A' represents A at a later time or something such as that. A being is identified by its four causes; since the cause of A' is D while D was not the cause of A in this scenario, A' and A should be identified as different objects, and therefore this circumstance is just a linear sequence which happens to have a certain cyclic similarity between the beings.

For temporal sequences, point 7 is strongly suggested by the second law of thermodynamics: if an equilibrium state exists for the system, then it cannot have stretched infinitely into the past because by now everything would be in equilibrium. Since we observe that the universe is not in equilibrium, and there is an known equilibrium state for the universe (an expanding, heat death scenario), it follows that the universe cannot extend infinitely into the past (whether temporally or in the chain of causality). What if we had an sequence of eternal inflation, where new pocket universes were springing into being, or the universe collapsed into a big crunch and then bounced to become another universe? This ideas are purely speculative, but even so we still cannot avoid the second law; the series of bounces would eventually hit a state where the universe would expand to a heat death, and after this it would bounce no more. The pocket universe generating multi-verse would hit a state where no more universes are generated, and that would also be that. If the system had extended infinitely back into the past, then we would be in that equilibrium state, and we are not. Therefore the system cannot have extended infinitely backwards, and the series of causes must terminate. Note also that the earlier derivation of the second law depends only on the assumptions of ergodicity and locality (that an effect only depends directly on its cause), which are not satisfied only in physics but also in any scenario beyond physics where the laws of causality hold and the range of states of possible beings is complete; any system of complex causes and effects must eventually tend towards some equilibrium.

One can raise an objection here. The second law is a physical principle. However, I am attempting to extend the argument beyond the material world into the immaterial world. As soon as we pass from physics into metaphysics, an argument based solely on physics ceases to be valid.

However, there is more to this than just the argument from thermodynamics. The sequence of causes must form either an accidental or instrumental series. And it must remain as such even if we move beyond physics into some non-physical causes. Physics by itself just describes the nature of the causes. We then have to figure out whether the sequence is accidental or instrumental, which means invoking philosophical arguments. The philosophy tells us what to look for, and then the physics can tell us whether those conditions

are satisfied. My argument is that the second law of thermodynamics allows us to establish this as an instrumental sequence. Obviously while we are traversing the physical causes, then we can continue to use a physical argument. But eventually this sequence of physical causes must finish. There are then two possibilities. Either that last physical cause is the true end of the series. Or the sequence continues into some immaterial cause. I do not want to beg the question at this stage by asking whether or not there can be such immaterial causes. In the argument so far, it has not yet been proved that there can, but neither has it been proved that there can't. So best practice is to leave the question open until we see where the analysis takes us. Obviously the link between the immaterial cause and the first physical member of the chain is not something that can be described by physics, since physics is a description of the interactions between physical objects. It doesn't say that there cannot be immaterial causes, and nor does it say that it can't. The topic is just beyond the scope of physical theory. But once we have established the chain of physical causes as instrumental, which is one part of the complete chain if it continues into the immaterial, then the link between the immaterial and the first physical cause, whatever the precise details of the interaction happen to be, must also be instrumental. Which means that the sequence must eventually terminate. One cannot switch from an accidental series of causes to an instrumental series midway through the sequence. So if we establish that the first part of the series is instrumental, then the philosophical argument can take over. This allows us to pass through the barrier between the material and immaterial. Of course, this doesn't prove that the first cause is immaterial: that still has to be demonstrated when we consider the nature of the first cause. It might be that, for all we know at this point, that the sequence of causes ends with some lump of matter. If we relied solely on an argument for physics, then we wouldn't be able to go beyond that point, whether or not it was the true end point of the sequence. But by reflecting on the philosophical nature of the chain of causes, we remove that restriction from our argument.

So is this an instrumental series? It is worth reviewing the four criteria highlighted in the previous chapter (page 405).

- **Something is passed down the chain of causes.** In this case, at least for the material part of the sequence, that is energy. A photon can only be emitted if the electron was in an excited state with sufficient energy, and that energy is transferred to the photon, which is then transferred to the next particle in the chain, and so on.
- **The objects cannot generate this by themselves.** Energy cannot be spontaneous generated by any physical particle, but only transferred from one particle to another. This is guaranteed by the conservation of energy, which is ultimately derived from locality.
- **The chain needs a direction.** This is guaranteed by the second law of thermodynamics and the observation that the universe is not in statistical equilibrium, but there is a state of statistical equilibrium to which the universe is tending.
- **Each interaction needs to be local.** This is guaranteed by the locality of the Lagrangian.

So at least the physical part of the sequence of causes is instrumental. But this means that the whole sequence of causes is instrumental, whether or not it terminates with something material or immaterial. If the series is instrumental, it must have a first member: something that can initiate the motion without having been moved itself by something previous to it in the chain of causes. The only question is what is the nature of this first cause; which will be discussed below.

Alternatively, one can reason from the mathematical impossibility of a past infinity of countable objects (Spitzer, 2010), since each transition between cause and effect can be counted as one more step in the chain. In mathematics, the concept of infinity is defined as an idealised limit, but as such it is not something one can encounter in practice. For

example, while one often writes the sum

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 2,$$

this in reality represents only a limiting case: it states that as you add more terms to the sequence, you asymptotically approach 2. For example, if one takes a sequence, and adds one more member to it, and keeps doing the same, then one would never reach infinity. What if one continued the process for ever? But 'for ever' just shifts the same objection into a different context; one could keep counting back one second, but one never reaches an infinite number of past seconds. If there were an infinite number of past causes, then there would be the same number of past causes for two successive members of the chain of causality, which is a contradiction; the contradiction lies in the concept of a countable infinity. Thus the chain of causes cannot continue indefinitely.

Point 8 thus follows from the previous arguments. If something exists contingently, then there must also be something that exists necessarily. Point 9 removes that 'if'. This point cannot be proved by logic, but is obviously correct from observation. We are thus left with the conclusion: that there must be something which exists necessarily. □

Objection:
Objections to this cosmological argument usually attack one of the premises. The main attack of the enlightenment philosophers was on the principle of causality, saying that the scholastic notion of causality is no longer valid. •

Response:
The attacks on causality were based on an outdated physics. As shown earlier, modern physics is fully compatible with the scholastic view of causality, and inconsistent with views of causality developed in the enlightenment. For example, Hume and Kant both based their objections on what we now know to be a faulty understanding of causality. ○

Objection:
The cosmological argument is ultimately dependent on the ontological argument, which is needed to link it to the definition of God. Thus as an argument for God, this is based entirely on arbitrary definitions and wordings and not on observation and thus more a matter of word play rather than fact. •

Response:
This objection is simply false, and arises from using the ontological argument rather than the cosmological argument to define what is meant by 'God'. The objection assumes that I am trying to use this argument to reach a *logically necessary being*; but the argument is not intended to do that, terminating only with the weaker condition of a *necessary being*. It is true that the cosmological argument doesn't obviously imply that its termination is the greatest possible being, but it doesn't need to because that is not part of the definition of God (we might - or might not - later conclude that God is the greatest possible being, but this objection relies on God being defined in that way). I have defended points 2 and 5 on the basis of arguments established from physics, which are based on empirical evidence as well as rational arguments. But point 9 comes entirely from observation (I do not think that it even requires the assumption that our senses are reliable); it cannot be proved rationally. Since an ontological argument is one that arises from reason alone, and this argument does

not rely on pure reason, it is not ontological. ○

Objection:
Alternatively, one may attack the statement that there cannot be an infinite chain in the past, saying that there is nothing unimaginable about an infinite chain of causes, for example by extrapolating from a finite chain. •

Response:
Once again, we need to distinguish between what we can imagine and what reason shows us, and in any case we cannot imagine nor conceive of an infinite chain: infinity is just too big, so even if we could rely on imagination or intuition it would fail us. Arguments of this nature usually start by saying 'we can always add one more cause, so just continue that process for ever'. However, this smuggles in the idea that an infinite series is possible into the premises by use the term 'for ever'; the question is wholly whether that 'for ever' is possible in principle or even makes any sense. You might ask 'Why not?', but the point is that our experience is only for limited amounts of time; the difference between finite time and infinite time is not one of degree, but of kind. Infinity is wholly different to any finite number. Thus extrapolating from our experience of finite numbers to infinity is flawed. In mathematics, infinity is usually treated in terms of a limiting procedure: in continuous mathematics, we ask what does this integral from zero to $1/\delta$ approach as $\delta \to 0$; or in discrete mathematics 'What does this series approach as n becomes larger?' So, for example, in a geometric series $1+a+a^2+a^3+a^4+\ldots$ we consider the result for a finite number n terms $(1-a^n)/(1-a)$, and note that for $|a| < 1$ as n increases this approaches the value $1/(1-a)$; the result is a limiting value. In continuous mathematics $a^n \to 0$ in the limit that $1/n \to 0$, and then continue this statement to the case of increasing discrete n.[1]. ∞ in mathematics is just a useful shorthand for this sort of limit, in both continuous and discrete arithmetic. The proof for the sum of a series just applies to a finite series, since nobody has nor can in principle write down a series with an infinite number of terms.

Most attacks on the idea of the impossibility of an infinite series fail to make the distinction between an accidentally ordered and essentially ordered series, and thus fail. It is not enough to assert that an infinite series is possible without showing it for the particular type of infinite series involved in the cosmological argument; and, of course to find flaws in all the various arguments against such a series. For example, one would have to show that the second law of thermodynamics doesn't hold; and that the chain of causality cannot be thought of as a Markov chain. Good luck with that. And that is before we consider the alternative arguments, such as from the impossibility of a *essentially ordered series* not terminating, or the impossibility of a countable infinity. I have not seen any strong argument against the best types of the argument against an infinite series of efficient causes; usually what are attacked are various straw men. ○

Objection:
If God created the universe, then who created God? •

Response:
It should by now be obvious why this argument is treated with derision by knowledgeable

[1]One can also replace a discrete sum as an integral in the continuous complex plane over a function multiplied by $\cot \pi n$ as another way of writing the discrete infinity in terms of the continuous concept of a limit.

theologians. Firstly, it is usually accompanied by a misunderstanding of the principle of causality, as though it were 'everything has a cause' rather than, as correctly stated, 'everything contingent has a cause' (conjecture 13.1.8). The first premise of the argument above was (in effect) 'Either something has a cause or it doesn't', which is not in contradiction to the conclusion that there is something without a cause. Only Hume's misrepresentation of the cosmological argument contains that implicit contradiction so beloved of atheists from Mill, Bertrand Russell and onwards.

Secondly, we must ask what is meant by 'God' when this question is asked, and this concept is often not precisely defined in those atheist writings which deploy this argument. But since the question assumes that 'God' is both a termination of the chain of causality and contingent (otherwise it makes no sense to ask what created Him), we see that this question is trying to argue against the idea of a contingent first cause. Fine: I also opposed a contingent first cause while supporting point 5 above. In other words, once we have untangled the definition of 'God' used in question – which is not the same definition that theists use – we see that this question does not provide an argument against theism, rather it provides part of an argument in support of it. Atheists would do better to try to show that there could be a contingent first cause (which would not be the God of theism) rather than to argue against it.

Alternatively, we could substitute the definition of God actually used by theists into the question, which (in part) states that God is a necessary being. Since a necessary being is one that cannot in principle be caused, we can rephrase the question as 'What caused a being that by definition could not have been caused?' But this question implies a self-contradictory principle; the question cannot be answered, not because the idea of God is incoherent, but because the question itself is incoherent. ○

Objection:
Finally, there are those who look at the conclusion of the cosmological argument, and say, 'So what? You have not proved the existence of God[2]. What about the various properties you give of God such as omnipotence, being personal and so on?' •

Response:
If I ended this work here, or (for example) Thomas Aquinas ended his *Summa Theologica* after the second question, then this objection would have some force. But neither I nor the greater thinkers who proceeded me have done so. Firstly, we have not yet reached the theist's God, but only taken the first step in a longer journey; secondly we must also discern what properties necessarily follow from the being whose definition we will give below, and we shall then see why we may rightly regard this being as God. The cosmological argument leads us to a definition of God, and it can be shown that such a God must possess the divine attributes, but the cosmological argument itself does not demonstrate this last step. But then no theist has ever claimed that it does. That requires further argumentation, which is subsequently given. But every theist of note in their apologetic work will provide that argumentation in the chapters following the cosmological argument, or will refer back to another writer who has done so. It is only those atheists silly enough to make this objection who are so sloppy in their thinking they will pull out the cosmological argument from its wider context, and then assert that the follow up work was never done without ever reading the writers they are criticising to see if that assertion is true. ○

[2]For example Hume, *Dialogues concerning natural religion.*

CONJECTURE 13.1.10. *There is at least one necessary being that is self-existent*

Proof:

(1) Everything that exists necessarily either derives its existence from something else or from itself.
(2) If it derives its existence from something else, that exists necessarily.
(3) There is therefore a chain of causes which either terminates or continues indefinitely.
(4) It cannot continue indefinitely.
(5) It cannot terminate with something that derives its existence from another.
(6) It must terminate with something whose existence is contained within itself.
(7) There is at least one being which exists necessarily.
(8) Therefore, there is at least one being which derives its existence from itself.

■

Discussion:
By derive its existence, I am asking what is the reason that it exists. Deriving its existence from itself (self-existent) means that there is nothing exterior to it that causes its existence. Point 1 follows obviously, as the two options are mutually exclusive and complete. Point 2 requires a little more explanation. An analogy is the music played by a musician - if the music exists necessarily, then it is clear that the musician must always be playing it, which means that the musician exists necessarily. More generally, we can see that if there is no time when a necessary being comes into existence, then it cannot have an instantaneous efficient cause. Another analogy is the hydrogen atom, which can only exist if its constituent electron and proton exist. It must therefore derive its existence in some other way, and the only option available to us (as long as it is not self-existent) is a continual generation, in this case the electromagnetic field which binds the proton and electron together. It exists because there is something else that continually acts to keep it in existence. For the necessary being, if the process were not continual, but stopped at some point, then from that point there would in effect be an instantaneous efficient cause, which is impossible for a necessary being. If something else continually causes it to exist necessarily, then that something else must also exist necessarily. Point 3 then follows; the cause of our original necessary being must either have a continuous cause, or be self existent. This leads to a chain, which either continues indefinitely or terminates with some self-existent being.

The infinite chain, however, is impossible because this is a sequence of instrumental causes, which must have an end. A sequence of instrumental causes $A \to B \to C \to D$ can always be collapsed into two members, the first and the last, $A \to D$ with B and C as intermediate steps. But this implies that there is a first member of the chain. Alternatively, we can argue that if the chain of explanation continues indefinitely then there is ultimately no explanation for existence; existence is in principle (and not just in practice) unexplainable. This, however, violates our assumption that the universe is rational and understandable. Furthermore, the explanations in this chain are are of the form 'A can only exist if B exists', 'B can only exist if C exists.' If the chain terminates at some point with something that is self-existent, then we find 'Z exists', which provides justification for our observation or inference from observation that 'A exists.' If, however, the chain is infinite then everything only remains an 'if'. We have no rational grounds for stating that A can exist even though we observe that it does: a contradiction. Thus, we conclude that the infinite chain is impossible.

Denying point 5 leads directly to a contradiction; the termination of the chain must be with something that needs nothing outside itself to explain its existence. We are thus left with point 6. Point 7 was proved above (conjecture 13.1.9), and the conclusion follows

directly from 6 and 7. □

CONJECTURE 13.1.11. *For the being shown to exist in conjecture 13.1.10, its existence is essential*

Proof:

(1) The essence of a being is that which distinguishes it from other substances and which always applies to it as long as it exists.
(2) A being whose existence is essential to it must always exist.
(3) A being whose existence is essential to it cannot derive its existence from something else.
(4) A being whose existence is essential to it must be self-existent.
(5) A being that is self-existent must be essentially existent.
(6) Therefore, the essence of the totality of self-existent beings is their existence.

■

Discussion:
Point 1 is a paraphrase of the definition of essence. A contingent being's essence cannot include the property of existence. The form of a dog, for example, is well defined whether or not the dog exists (for example, we may remember a dog which has recently died). This implies that existence is not essential to the dog, as it may or may not exist. A being which exists essentially must therefore be necessary. This proves point 2. Similarly, a derived existence A cannot exist essentially, because by definition 'A can only exist if B exists' implies that A may or may not exist depending on the existence of something else. Only something which is self existent and necessary could be said to exist essentially. This demonstrates point 4. But then this point, that it could be said to exist essentially, distinguishes the self-existent being from everything else and always applies to it.

The defining attribute of such a being is therefore its mode of existence, and thus it can be identified by the statement that *It is*, or (if It referred to itself in first person) *I am.* □

Objection:
Am I treating existence as a property? •

Response:
As discussed in my response to Kant in section 12.4.2, it is quite reasonable in an Aristotelian framework to treat existence as a property. ○

DEFINITION 13.1.12. **God:**
We shall see shortly that there can only be one necessary self-existent being. We define the word *God* to refer to this being. △

Objection:
This definition is very different from that used in traditional religions, such as Christianity

and Islam. •

Response:

> Then Moses said to God, 'If I come to the people of Israel and say to them, "The God of your fathers has sent me to you," and they ask me, "What is his name? what shall I say to them?" God said to Moses, 'I am who *I am*.' And he said, 'Say this to the people of Israel,"*I am* has sent me to you." '[3]

Equally,

> For as the Father has life in himself, so he has granted the Son also to have life in himself[4].

The Jewish and Christian God (if He exists), claims to be God as I have defined it here. If this claim is true, He may be more than what is implied by this definition, but He is certainly not less than it. It is therefore appropriate to use the word *God* to both describe this metaphysical concept and the being known as *I am* (or Yahweh/Jehovah/El) in the Bible. o

Discussion:

Most of what are called gods are not Gods – they are more like angels or demons; their existence or non-existence is irrelevant to the claims of theism It should be noted that many of those called gods do not satisfy this definition; for example the gods of the Greek, Norse and Hindu pantheons are not necessary, self-existent beings. For example, we are told that Apollo was born of Zeus, Zeus was the son of the Titans, who themselves emerged from the primeval chaos. Such beings are closer to angels or demons in Christian theology (and most likely demons, since they lead worshippers away from God) than gods. The closest thing in Greek mythology to the God I am describing here is the primeval chaos from which everything else sprang.

This means that those who claim that they just disbelieve 'in one God more' are making several logical fallacies: a category error, an error in fact, and the evidential error. The first is to lumping the likes of Zeus, Odin, or the Hindu or Chinese Pantheon as though they were the same type of being as God: they are not. These gods are of a fundamentally different kind from the God that is the originator of everything, self-existent and necessary. Secondly, it is not necessary to disbelieve in such beings to be a theist. If Zeus exists (in the category that the Greeks called gods and I call demons), it makes no difference to the theist's belief. Theists are therefore rightly not atheists in regards to such 'gods' but agnostic[5]. I do not personally accept the existence of Zeus, but then I have not examined the evidence presented in his favour; I can hardly then claim to be expert enough to absolutely dismiss

---

[3]Exodus 3:14

[4]John 5:26. *Life in himself* is distinct from *life from someone else.* Life in John's gospel sometimes refers to spiritual life (an analogy often found in Christian literature including the Bible; it refers to being in a right relationship with God) and sometimes to physical life (what we usually think of as *life*). In the case of physical life, we obtain our life from our parents; our life is not in ourselves but from another. In this sense, having life in Himself is close to what I have termed self-existence (since, for a living organism, if it lives it exists and if it exists according to its proper form it has life). If the phrase refers to spiritual life, then for us our relationship with God is not something we can restore by ourselves; we can at best cooperate with and permit the work of God in ourself. Our spiritual life is thus derived from God. For God, again, if He exists then he must be spiritually alive (or else in a state of self-contradiction), if spiritually alive then he must exist. Thus again this passage implies God's self-existence, even if it is not obviously discussing it at first glance. Within the context of the doctrine of the Trinity (discussed later), there is no difficulty in saying that both the Father and Son are self-existent while maintaining that there is only one self-existent being, as the Father and the Son are of one substance.

[5]See, for example, 1 Corinthians 10:20, where Paul relates participation in pagan sacrifices to the Greek gods (for example by knowingly eating meat dedicated to the idols) with demon-worship.

it. But such evidence is irrelevant to me, since my beliefs and manner of life will not be modified whatever it claims: even if true, it is as relevant to my conclusions about God as much as the possible existence of an Indian called Mahatma Tendulkar. On the other hand, the God which I am discussing here, if He exists, is of the utmost importance to all of us, as we shall see. Thirdly, the degree of evidence for the one God of the monotheistic religions and the many gods of polytheism is entirely different. Metaphysically, we shall see that if our premises and reasoning are correct God must exist for anything to exist; this cannot be said for any believer in Zeus, who must rely on observational evidence. Similarly, the observational evidence in support of Christianity is far more reliable than that provided by the various Greek, Norse or other polytheist myths.

Also note that this definition of God is not (however bizarre the conjunction of words may seem) incompatible with atheism. One form of atheist may reject the idea of a necessary, self-existent being; I believe that these people are entirely mistaken. Another group of atheists may affirm the existence of a necessary, self-existent being, but deny that this God has the properties usually attributed to Him by theists and deists, most importantly that God is an immaterial mind. I will later use the expression 'atheistic God' to denote a necessary, self existent being which cannot be though of as a mind. While the first of these atheistic positions is, I think, wholly irrational, the second is certainly more plausible. I have certainly not disproved every form of atheism in the discussion so far. □

Comment:
Note that at this stage of the argument, we have not yet shown that God has the usual properties assigned to Him; that He is an immaterial mind; that He is omnipotent; omniscient; omnipresent, benevolent and so on. These must be regarded as open questions until we either demonstrate that God has these properties or demonstrate that He does not. Readers may question whether I am free to define God in such a way that does not obviously (if at all) possess these properties: however, I believe that I am justified in doing so. The standard picture of God is also emphatic that He is self-existent. I must therefore choose one and only one of these properties as the basis of my definition, and then examine whether the others also apply to this being (to choose more than one of these would be begging the question of whether they both apply to the same being). Since self-existence is the first of these properties we encounter, it is convenient to use it as the basis of my definition. (For the use of the pronoun 'He' to describe God, I must beg the indulgence of the reader; I have, of course, taken it from Christian tradition rather than the arguments presented in this chapter. It is ultimately derived from the Christian views of Masculinity and Femininity, and in particular that they are primarily distinguished by their proper relationship to each other; calling 'God' Father implies a very different character to God and relationship between God and ourselves than if God were called 'Mother', as everyone who paid attention while being raised by both a mother and father, in a well-functioning household, would testify. There are many religions which worship a mother god; their gods are all very different from the Christian God.) □

13.2. The Attributes of God

DEFINITION 13.2.1. **Immutable:**
Immutable means incapable of changing. △

CONJECTURE 13.2.2. *God is immutable*

Proof:

(1) A being that changes ceases to be what it is actually and becomes what it could be potentially.
(2) However, a necessary being cannot by definition cease to be what it is.
(3) Therefore, a self-existent necessary being cannot change, and is immutable.

■

Discussion:
The first point is the definition of what it means to change, the second follows from the definition of a necessary being. Thus the conclusion follows directly from the premises and the definition of God (13.1.12). □

Objection:
If the necessary being consists of several potentia, then what it is is determined by its form (the set of all potentia), and change from one potentia to another is possible without the being itself coming into existence. •

Response:
God was defined in terms of the terminus of the chain of efficient causality. Each efficient cause is an actual object, and consequently exists in one particular potentia when acting as an efficient cause. If God continually moved from one potentia to another, then none of these could be considered the terminus of the chain of causality. We would be left with an infinite series of efficient causes, which was proved impossible above. ○

Objection:
Verses such as Exodus 32:14 ('And the Lord relented from the disaster that he had spoken of bringing on his people') and Jeremiah 18:9 ('And if at any time I declare concerning a nation or a kingdom that I will build and plant it, and if it does evil in my sight, not listening to my voice, then I will relent of the good that I had intended to do to it.') show the Christian God changing His mind and purposes. Therefore the Christian God is not immutable, and thus is not the self-existent necessary being described above. •

Response:
On the contrary,

> God is not man, that he should lie, or a son of man, that he should change his mind. Has he said, and will he not do it? Or has he spoken, and will he not fulfil it[6]?

and

> For I the Lord do not change[7].

Furthermore,

---

[6]Numbers 6:23.
[7]Malachi 3:6.

> Every good gift and every perfect gift is from above, coming down from the Father of lights with whom there is no variation or shadow due to change[8].

Throughout all the Bible, God's purposes remain constant: that he should deal justly, disciplining those who err, and having mercy on those who earnestly desire righteousness. God's actions within time may seem to change, but that is because they deal with changing circumstances. It is not God which changes, but the situation in the world. Just as a virtuous man must adapt to the situation to remain true to his character, so God acts in different ways in different situations to be true to His unchanging character; just as the desire to help an old lady cross the street may be either virtuous or vicious depending on which side of the street she wants to be. It is thus not God that changes when He relents of an action, but the people who responded to Him either with repentance, bringing blessing, or with arrogance, bringing upon themselves a curse. Indeed, in each of the passages where God is said to relent of some destructive punishment, there is a change in circumstance: for example, the people turn away from their evil, repent, and so it would no longer be just to punish them nor necessary to discipline them. The use of the word 'relent' in this context should thus be treated metaphorically. ○

CONJECTURE 13.2.3. *God is a being of pure actuality.*

Proof:

(1) Unactualised potentia express a capacity to change while remaining of the same substance.
(2) God is incapable of change.
(3) Therefore God has no unactualised potentia.
(4) God is actual.
(5) Therefore God is a being of pure actuality.

■

Discussion:
Point 1 is from the definition of potentia (definition 4.5.2), while point 2 was shown above (conjecture 13.2.2). Given that God exists, then He is actual since everything that exists has one actualised potentia. The definition of a being of pure actuality is that its form is not a mixture of actuality and potentiality, i.e. its only possible state is actual. From what proceeds it is clear that God is a being of pure actuality. □

DEFINITION 13.2.4. **Part**:
A part of a being is an aspect of a being that can be taken out of the whole and studied in isolation. This does not mean that its proper functioning would not require the rest of the being, but only that it is not incoherent to consider it in isolation. We can consider physical parts, such as the heart, lungs, or legs of a man, or metaphysical parts, such as a being's form or matter. One can take the heart out of a man, and it will maintain its existence, at least for a short time. As another example, the bare (unrenormalised) electron is not a part of an atom, because it cannot be separated from interactions with its own electromagnetic field, and neither can that electromagnetic field in isolation from the electron; the renormalised

[8] James 1:17.

electron, which considers the bare electron together with this self-interaction can be considered a part of an atom since it can have an independent existence apart from the atom. △

Discussion:
Some (not necessarily or even usually all) properties of the whole are taken from the properties of its component parts. In particular, if an essential part of a being is contingent then the whole must also be contingent. This implies that if the whole is necessary, then each part must also be necessary. □

DEFINITION 13.2.5. **Simple**:
A simple being is not composed of parts, whether physical or metaphysical. Note that a material being cannot be simple, as it is the union of form and matter. A simple being is necessarily of one substance. △

DEFINITION 13.2.6. **Composite**:
A composite being is one that is not simple. △

CONJECTURE 13.2.7. *God is essentially simple*

Proof:

(1) If God were composed of parts, then each part in itself must be a necessary self-existent being.
(2) In a composite self-existent being, there must be a mixture of potentiality and actuality.
(3) God is pure actuality without potentiality.
(4) Therefore God cannot be composite.
(5) Therefore God is simple.

■

Discussion:
Point 1 is not intuitively obvious: after all, a hydrogen atom is composed of parts, but these parts are not essentially hydrogen. However, in this case, we see that if A, a supposed part of God, were not self-existent, then it must be of derived existence. If so, there must be something, G, prior to A and different to A which is self-existent. By definition, G would then be God (or a God); however this means that A could not be even a part of God as it is different in essence from G, contradicting our original statement. Therefore each part of God, to be God, must be self-existent.

Now suppose that God were both composed of parts A and B and necessary. From the definition of 'part', these parts must in some sense be independent of each other, so that one can exist without another. There may or may not also be some unifying principle binding A to B; every composite being we are aware of has such a unifying principle. However, God cannot have such a principle. Such a bond would not be necessary, because if it was then A and B would not be independent of each other and thus parts, and it must be part of God, because otherwise God's nature would depend on something outside Himself, which contradicts God's self-existence. But it cannot be part of God, because every part of God

would in itself be necessary and self-existent, and such a bond cannot be necessary, as already shown. Therefore, if God is composed of parts, those parts cannot be in any way joined together or inter-related; He would just be a collection of distinct objects. For example, we can have that A can exist independently without B, and B can exist independently without A, and God would be the combination between A and B. In this case, there are several different states in which God could exist, A, B, or various different combinations between A and B. These states, however, would represent potentia. If God were 50% A and 50% B (for example), there must be some reason to explain why God would take these proportions, and this reason would be ontologically prior to God and thus inconsistent with God's self-existence. Thus a composite self-existent being being must contain potentia, which demonstrates point 2. However, above we saw that God lacked potentia, hence God is not composite, and since everything that is not composite is simple, I conclude that God is simple.

Note, however, that this argument only applies to God's essence or substance. In this conjecture, we have shown that God is of one substance (i.e. simple and devoid of parts), but this does not show that there can be no distinctions in God which are distinctions in something other than God's essence, for example distinctions in relationship, where both the source and result of that relationship are identical in essence (i.e. both self-existent), but distinct in how they relate to each other. If the relational bond is necessary, then the two hypostasis (as we may call them) cannot be separated from each other, and they cannot rightly be called parts, but they are nonetheless distinct.

This idea of the possibility of distinction within a simple substance answers an important objection to God's simplicity: if God is an immaterial mind (as I shall show later), then how could He be absolutely simple with neither difference nor distinction within him? Thought, after all, surely requires some degree of complexity. But this is not valid if we allow for distinction within God while (rightly) repudiating the error of saying that there is difference; these distinctions are enough to permit thought. For example, there is a clear distinction between a thinker and the thought, though this is not one of essence, and nor are they parts of an essence because they cannot be isolated from each other: one cannot have a thinker without a thought, and nor can one have a thought without a thinker. Thus the distinction between thinker and thought is primarily a matter of relation rather than substance, and it is possible for an immaterial mind to be simple in substance. Indeed, an immaterial mind must have hypostasis distinguished by their relationship to each other.

Also note that divine simplicity means that the physical universe itself cannot be God; since the universe, being material, must be composite. God must be immaterial (this is elaborated on below).

While many Christian theologians today question divine simplicity, it is a key element of the classical understanding of God and I do not think that one can reasonably deny it while remaining consistent with Christian revelation and tradition. It is a key element of both the Platonic (early Church and Eastern Orthodox) and Aristotelian (Roman Catholic and magisterial Protestant) expressions of Christianity, and contained in most Protestant Confessions. The doctrine of the Trinity is not contrary to it: obviously the people who formalised that doctrine were classical theists (of the Platonic variety) who held to the doctrine of divine simplicity – much of the problem in formulating the doctrine of the Trinity was to make it consistent with divine simplicity, and because the requirement of simplicity eliminates other options, we can be sure that if the basic tenets of theism are correct then the classical statement of the doctrine of the Trinity is correct[9]. □

[9]See (Sanlon, 2014) and (Dolezal, 2011) for a defence of divine simplicity.

Objection:
God, as viewed by most Christians, is seen as essentially having many distinct attributes, such as omnipotence, omniscience, benevolence, justice, mercy, agape, and so on. Thus there are many boundaries within God separating these attributes, since they describe entirely different quantities. However, a simple being cannot have boundaries within it. Therefore the God of Christianity cannot be simple. Therefore the necessary self-existent being described above cannot be the God of Christianity. •

Response:
This objection has one loophole: it requires that omnipotence, omniscience, and so on are distinct within God. If, however, they are entirely equivalent to each other, then God does not have many attributes, merely one attribute which manifests itself in different ways. This would be allowed if, for example, defining God is 'the omnipotent being' is entirely equivalent to defining God as the 'necessary self-existent being' or 'the omniscient being', and so on for all the attributes assigned to God. In this case, the simplicity of God is maintained since the attributes are not distinct but the same thing described in different ways. I have not yet proved that this is the case, but nor can it be disproved from what I have discussed so far.

The obvious response to this answer would be to say 'Omnipotence is to power what omniscience is to knowledge; but knowledge and power are not equivalent. Therefore omnipotence and omniscience cannot be equivalent.' However, this response is making the mistake of viewing power in us and power in God as being the same thing, differing only in degree; our knowledge and God's knowledge also differing only in degree but not in kind. It is not clear that this is the case. Rather, we apply knowledge, power and so on to God in a way that is analogous to how they are applied to us; God's knowledge plays the same role for Him as our knowledge does for us, but is of a very different nature, as shall be clear when we later discuss what it means for God to know something. Therefore this response does not hold up.

The objection therefore rests on whether God (as defined in this discussion) must have these attributes and any being which has these attributes must be God, or whether these attributes would be an addition to His essential nature. The appropriate place to discuss this, however, is when we discuss each of God's attributes.

The important thing to always remember is that the difference between God and ourselves and the things we observe around us is not only one of degree but of kind. The Greek or Norse gods, for example, were seen as beings like us, but only better. But the theistic version of God is entirely different from us. We cannot use our own experiences and extrapolate them upwards to God because there is a complete break between the two. This, at least, is what all major Christian theologians have held (until, perhaps, the modern era); and all have held divine simplicity to be an essential part of our understanding of God. (And this is, of course true, for all theists, and not just Christians.)

So what of the statement in Genesis that mankind is made in God's image and likeness? Doesn't that show a certain similarity between us and God? Resemblance, yes, but similarity no. We often think of images today as in reflections in a mirror, but an example closer to the author's intent in the Genesis text would be a statue. If a skilled artisan made a statue of (say) Cicero, then that statue would be made in the image of Cicero. There would, of course, be resemblance between the external appearance of the statue and of the man, if the artist is good enough. But that still doesn't change the fact that the statue is fundamentally different from Cicero. You can't say that Cicero is like the statue, only more so, for he possessed numerous additional powers and qualities that statues lack. So why can't we say that Cicero is the image of the statue? Two ideas occur to me: firstly, Cicero came first: obviously first chronologically, but for my purposes the more important sense is first in order of inspiration. Secondly, we can think about our language. When we say that a

statue resembles Cicero, then we mean that it has a close physical similarity: we think of those things that they have in common. When we say that Cicero resembles a statue, then we think of it as a negation of their differences: we mean that Cicero is standing immobile, and erect. The natural purpose of the statue is only to bear a close physical resemblance to the thing it depicts for the sake of art. It is a medium that focuses our mind onto the thing that it depicts. To say that a statue is a good resemblance is to say that it does this job well. However, the natural purposes of a man, those things that distinguish him, are very different. They include the tendency to receive nourishment, the tendency towards self-movement, the tendency towards thinking, and so on. So when we say that a man resembles a statue, we think in terms of these purposes, and see him as cold, motionless, and dumb. So it is when we say that mankind is made in the image of God. Can we say that the statue resembles a part of Cicero, namely his physical appearance? No, because Cicero's appearance cannot be separated from Cicero the man. It is an aspect of who he is (baring that, for Cicero, unlike God, His appearance is subject to change and decay). We cannot separate Cicero's appearance from him except as a part of a mental simplification which we will eventually have to undo if we want a complete understanding of Cicero.

We can also say that a written description of Cicero is made in the image of Cicero. Yet, although the writing is very different from the statue, they both refer back to the same unified individual. They both reflect Cicero's nature, but different aspects of it (for example, the writing can describe the motion of Cicero's lips when he smiled, which a statue cannot; while the statue can convey the precise shape of his nose in a way that the written description cannot), but the differences between the reflections are not due to divisions in the man they describe, but because the two different images are captured in different media.

So when we say that man is made in the image of God, we mean it in the sense that the statue is made in the image of Cicero. When we describe God as having power, or knowledge, we do not mean that there is an attribute of God that is some way resembles a part of us (or something we are familiar with). Rather, we are saying that there is a part of us that in certain respects (though not every respect) resembles God, not a part of God, but God as a unity. Thus we should not extrapolate from what we mean by power or knowledge and attribute them to God, any more than we extrapolate from what it means for a statue to have a certain appearance to what it means for dear old Cicero to have that appearance, and suppose that the means by which that appearance came about or the matter that is shaped in that way are the same. In this way, we may use the same word 'knowledge' for both ourselves and God, but that does not mean that God's experience of possessing knowledge is the same as ours, nor that God's means of acquiring or retaining knowledge is the same as ours, or that the concept of knowledge in God is identical to the concept of knowledge in us; merely that the two have certain similarities in both what they are and what they are for (but also various differences).

It is common to use the language of analogy here. We can use a word either directly (for example, when we say that the ball is red and the postbox is red, we mean the same thing by red in both cases), or we can use the same word to refer to two different concepts (I used the bat to hit a ball; A bat flew out of the belfry at twilight; we do not use a flying mammal to hit the ball), or analogously. By analogously, we mean that there is some resemblance in some areas, but not in all areas. An analogy is a simplification that helps us to understand something which is harder to grasp or imagine. In the same way, when we discuss God's power or knowledge, we mean that there is some resemblance (but not a complete resemblance) between our power and God's nature. It is a picture which helps us to understand something about God; but it is only a picture, and once we have grasped the main point in question, it should be discarded lest the differences between it and God lead us to misunderstand God; just as when we see Cicero in person, we no longer need the statue to focus our minds on him.

For example, when we discuss wave particle duality, we say that there is something in a quantum field which is analogous to the classical particle and something within it which is analogous to the classical wave. The classical particle and classical wave are obviously not the same thing – indeed, they are contrary to each other; however this does not mean that the idea of the quantum field tries to unite contradictory principles. The whole point of analogy is that the two ideas are not the same; there are differences as well as similarities. The fundamental quantum particle is not an incoherence, nor combined of material parts, even though something that is both a classical wave and classical particle would be incoherent.

God is impossible for us to imagine. This is obvious: our imagination works on those things we have first observed, and perhaps combines or reshapes them in various ways. But everything we imagine is like something we see; after all our imagination is designed to help us in the material world. But God is unlike anything we have seen. Our imagination fails. That doesn't mean that God can't be (to a limited extent) understood. Our understanding is not the same as our imagination; we don't have to imagine something to understand it; indeed often our limited imagination can be a barrier to a complete understanding. If we cannot imagine what knowledge is to God, then that is more a (quite understandable, given what we are) failure in our imagination rather than a failure in the overall concept. (We are, of course, in the process of this argument, not in a position to understand the overall concept, because I have not yet discussed these 'attributes' of God.)

Thus the objection is built on the assumption that our experience of these attributes is the same as how the words apply to God. From this we argue that because our imperfect knowledge is not identical to our imperfect power which is not identical to our imperfect agape love, they cannot be identical in God. But the first assumption is false, and therefore the objection to God's simplicity fails. ○

CONJECTURE 13.2.8. *God is unique*

Proof:

(1) If there are two or more necessary self-existent beings distinguished in essence, then there is a single essence which describes the species containing all of them.
(2) This essence contains the property of self-existence.
(3) But we have seen above that each necessary self-existent being or species must be simple.
(4) This contradicts our assumption that it is composed of the composite of several beings.
(5) Therefore there can only be at most one necessary self-existent being.

■

Discussion:
The species described in point one is the total of all necessary self-existent beings. So, suppose that A, B, and C are all self-existent necessary beings and the only self-existent necessary beings. Then the species is the set $\{A, B, C\}$ which contains all of them. Since the description of this set distinguishes these beings from those which are not this beings, it is valid to describe it as the essence of this species. It is obvious that this species does not depend on anything outside itself for its existence, so point 2 follows. Point 3 requires a little more expansion. The proof above was for a single, necessary and self-existent being. However, exactly the same argument applies for a self-existent necessary species: it cannot be composite. Thus the set must have a single member, and we conclude that there is only one God.

The unity of God may also be inferred from the unity of physics; there is a single principle acting on physical objects, which implies a single first cause. □

CONJECTURE 13.2.9. *God is immaterial.*

Proof:

(1) Every material substance is a composition of form and matter.
(2) God is essentially simple, and therefore not a composition of anything.
(3) Therefore God is not material.
(4) Therefore God is immaterial.

■

Discussion:
The first point follows from the definition of substance (definition 4.3.5), and the simplicity of God was proved above (conjecture 13.2.7). Form and matter are clearly parts of the substance, since one can grasp the form by itself in the intellect, and thus separate it from the matter. The conclusion that God is not a material substance follows. Since immaterial means 'a substance that is not material' (or a substance that is not made of matter), this shows that God is immaterial. □

DEFINITION 13.2.10. **Timeless:**
To be timeless is to be outside of time. To be in time means that we sense the passing of time; we are aware of the present, remember the past, but can only guess at the future. Events are viewed in succession. To be timeless means that there is no distinction within that being's nature between past, present and future. Events from all times are viewed not in succession but together.

Einstein has taught us that space and time are not distinct but similar. Just as we view the outside of a table as a two dimensional surface, so we should see space and time as a four dimensional surface (albeit not a flat surface, but one containing many bumps, divots and ripples). We are only capable of seeing three dimensional slices of this surface (not three dimensional slices at the 'same time' – indeed, the concept of describing things as happening at the 'same time' is ambiguous and should not be used except in a loose sense, but within the light cone - the cylinder passing from us backwards in time and outwards in space). To be timeless means to be able to see this whole surface in its entirety in one 'instant' (where I am using the word *instant* in an metaphorical rather than literal sense). A timeless being is not constrained by the four dimensions of space and time, nor by the flow of time. This might be achieved by having the intellect manifest in higher dimensions, or by some other means – if it is inappropriate to explain such a being in terms of any dimensions of space and time.

The idea of seeing objects at different times together rather than in succession is familiar to us. For example, when I look out of my window, I see the church spire, the moon, and a star in the same glance. But I do not see them all as they appear at the same time; I see the church spire almost instantaneously, the moon as it was about a second ago, and the stars as they were several years, decades or centuries ago. Thus I see three different moments in time together. Indeed, with a combination of mirrors and refractions, we can in principle view any objects at any past time in our past light cone together. The difference

between this analogy and God is, of course, that God does not use light to observe, so is not restricted to the past light cone and can view our future together with our past.

The opposite of timeless is temporal - existing in time, aware of the progress of time; only able to perceive within the past light cone. Δ

CONJECTURE 13.2.11. *God is timeless*

Proof:

(1) Time and space began together with the universe.
(2) Time and space are not self-existent.
(3) The cause of time and space is not essentially part of time and space.
(4) God is the cause of time and space, either directly, or indirectly through intermediaries.
(5) Therefore God is not essentially part of time and space.
(6) The essence of God is immutable.
(7) Therefore God cannot become temporal after the creation of time.
(8) Therefore God is essentially timeless.

■

Discussion:
In support of premise 1 I note that current cosmology points to a beginning of the material universe (in fact, there are several theorems which mandate that the universe has a beginning). Space and time are to be understood as being described by a vector space related to the metric of General Relativity, which in turn, through Einstein's field equation, is related to matter. Matter begins at a singularity when the vector space is infinitesimally small, and it impossible through physics to explore what happened 'before' this singularity - indeed the very question does not make sense. Thus the big bang not only marks the beginning of matter but the beginning of space and time.

Premise 2 then follows; if space and time 'began' then they must be contingent; if contingent then they cannot be self-existent, as the only self-existent being is necessary. They must therefore derive their existence from something else; have a cause. This sequence of causality is obviously not temporal, but an atemporal sequence of explanation. Therefore the explanatory cause of space and time cannot exist within space and time. Equally, the explanation of the explanation of space and time cannot exist within space and time (otherwise we would end up with a circle), and by definition the sequence of explanatory causes must terminate with God. Therefore, God, at least 'before' (here to be understood as earlier in the sequence of explanation) the creation of time and space must lie outside time and space.

We can then ask whether God could change to become temporal after creation (as proposed in (Craig, 2001)). This violates the proposition above that God is by essence immutable: the act of creation cannot change the nature of God. This leaves us with the conclusion that God is essentially timeless.

Now it might be objected that God's way of thinking (if He has indeed something analogous to our thoughts, which I have not yet proved) would then be very different to our own. It is difficult for us to conceive of or imagine a timeless mind, or a timeless being of any sort. However, this is no real objection, for no Christian would say that God's thoughts resemble ours in every or any respect. We are made in the likeness of God, in the sense that we possess intellect, will and freedom, but this does not mean that our intellect is of the same kind as God's intellect; we should rather view the relationship between the two as

analogous rather than a matter of degree. We should not suppose that reality is restricted to what we can imagine, and if our intellect confirms a proposition which our imagination cannot comprehend, we should accept the limitations of our less-sure imagination before concluding that our reason is at fault.

Observe also that the doctrine of God's timelessness refutes one argument raised at the beginning of the modern era that the laws of physics must be deterministic. For there it was held that God could only foreknow by predicting which is certainly true for temporal beings. And perfect prediction is only possible if the laws are deterministic. But for a timeless God, it is wrong to say 'foreknowing', for that word implies the concept of the progress of time. Rather, God observes[10] all of existence at once, from the beginning of time until its end; and it makes no difference whether the laws of physics are deterministic or indeterminate, as there can be knowledge of what is for us the future without needing prediction. Instead, God's timelessness is required if we are to combine God's omniscience with the observed indeterminate universe. □

Objection:
In the Christian scriptures, there is a development of revelation: God gradually reveals Himself over time. God starts with Abraham, reveals more to Moses, and so on until the revelation is complete in the person of Christ. This is a sequence of development in time. God did not send the law to Abraham, but it was only those after Moses and before Christ who were aware of and subject to the law. Therefore, God acts with respect to time, therefore temporally. But if God acts temporally, then he must be aware of the passage of time. Therefore the Christian God is not timeless. Therefore the self-existent being described here is not the Christian God. •

Response:
To be timeless does not mean that one is not aware of time, or cannot appreciate the passage of time as perceived by temporal beings, but that there is no passage of time within oneself. Were God to make Himself known through a special revelation at various discrete points within space-time, then those points must correspond to events in space-time. For a temporal being, they would appear to be ordered in sequence: first the earliest one, then the next one next, and so on. The temporal being would only be aware of the special revelation in their past, so the sequence would seem to them like a gradual development of revelation. Indeed, if the aim was for the temporal beings to understand the concepts, the timeless being will present the revelation according to their needs, revealing the most basic ideas earlier in time than the more advanced, or by using the passage of time through foreshadowing to show His divinity. To God, however, the revelation to Abraham, to Moses, in Christ and those subsequent to Christ including in our future is 'simultaneous', with each event placed at the point in space and time where it has the desired impact. That God is atemporal does not mean that He cannot appreciate our temporal natures. Therefore, any interaction between a timeless God and temporal creatures must appear to the temporal beings as being in time. Thus God may seem to us to act temporally without being subject to time Himself; and the above objection is invalid. ○

[10]The word *observe* here is used metaphorically. God obviously doesn't passively observe things in the same way that we do. God's contact with the universe is primarily through his being its sustaining and actualising cause, and one plausible suggestion is that God's knowledge of the things in the universe arises through His being the one bringing everything else from potential into actual existence. God's knowledge arises from God's active role in the universe rather than as a passive observer.

Objection:
Every cause precedes or is at the same time as its effect. Yet God is the first cause. Therefore God is a cause. Therefore God either proceeds the effect or is at the same time at the effect. However, both these statements only apply to a temporal being. Therefore God is temporal, not timeless. •

Response:
Every physical cause is simultaneous with its effect in time, as we know from the locality of field theories. However, God is not physical. Thus we need not suppose that the same principles apply in the same way.

God's timelessness means, by definition, that He encounters all events together rather than in succession. The timelessness of God means not that He is in no time, but rather that He relates to every time in the same way. We can say, in a metaphorical sense, that God is 'simultaneous' with every event and substance. Therefore this is no objection to God being both a cause and timeless. ○

DEFINITION 13.2.12. **Eternal**:
A being can be said to be eternal if it exists through all time.

This is the meaning of eternal as I shall use it in this discussion. However, it should be observed that sometimes (including by myself in different contexts) the word *eternal* is used in a different sense, that is without end in time (though possibly with a beginning). In the following, I use it as denying either a beginning or end in time. △

CONJECTURE 13.2.13. *God is eternal*

Proof:

(1) If a being is not eternal, then it must have a beginning.
(2) A necessary being cannot have a beginning.
(3) God is defined as the self-existent, necessary being.
(4) Therefore God is eternal.

■

Discussion:
This argument follows directly from the definitions above (definitions 13.1.6, 13.1.12 and 13.2.12). Note that existing through all time does not preclude that it exists outside time as well (although it does not deny it either). □

DEFINITION 13.2.14. **Omnipresence**:
Am Omnipresent being exists throughout all space. △

CONJECTURE 13.2.15. *God is omnipresent*

Proof:

(1) If a being is not omnipresent, then it must have a boundary in space, where it exists on one side of that boundary, and has infinitesimal probability of existing on the other.
(2) A necessary being cannot have a boundary in space.
(3) God is defined as the self-existent, necessary being.
(4) Therefore God is omnipresent.

■

Discussion:
The argument is obviously (and unsurprisingly, given the relationships between time and space) the same as that for God's eternity. □

13.3. The Mind of God

CONJECTURE 13.3.1. *There exits at least one being which terminates the chain of final causality. This being is God.*

Proof:

(1) Every contingent being must be a final cause of that which was its efficient cause.
(2) This creates a chain of final causality, which must either terminate in a necessary being or continue indefinitely.
(3) It cannot continue indefinitely.
(4) Every necessary being of derived existence is a possible final cause of that from which its existence is derived.
(5) This creates a chain of final causality, the teleological chain, which either continues indefinitely or terminates with a necessary self-existent being.
(6) It cannot continue indefinitely.
(7) Therefore the chain of final causality terminates with a self-existent, necessary being.
(8) This being is that defined above as God.

■

Discussion:
The first point is taken from the definition of final causality. Points 2 and 3 take us through the chain of contingent beings – including material beings such as ourselves – until we reach a necessary being (the chain cannot terminate with a contingent being, since that would have an efficient cause and thus be the final cause of something else); we then proceed in points 4 and 5 to proceed through the chain of necessary beings, until we reach the self-existent necessary being, God.

This argument is basically a recapitulation of the cosmological arguments above; and the reasoning is exactly the same; therefore I need not repeat it in detail. The only new point is the first premise, which was established in the proceeding discussion on efficient and final causality. The difference is that whereas the cosmological argument was based on efficient cause (or its equivalent when considering derived existence), this teleological argument is based on final causality. As we saw above, final causality and efficient causality are linked: everywhere where there is efficient causality there is also final causality in the

opposite direction. So if A is the efficient cause of B, B is a final cause of A. However, final and efficient causality are not equivalent because it does not follow that every final cause of A becomes actual; only one of the possible final causes becomes actual. So if we start with a being A with final causes B_1, B_2, B_3 and so on; and these have final causes $B_1 \to C_{11}, C_{12}, C_{13}$; $B_2 \to C_{21}, C_{22}, C_{23}$ then given C_{11} we conclude that there must have been a B_1 whose final causes include C_{11} and if there is a B_1 there must be an A whose final causes must include B_1. We cannot say that given A there must later be a C_{11}, only that C_{11} is possible. I should mention that the series of final causes cannot continue indefinitely is best shown by realising that it is instrumental in nature (since C_{11} can also be viewed as an indirect final cause of A, with B_1 merely an instrument in that process), and noting that every instrumental series must have a first member.

This argument then is related to the proof from existence and efficient causality above, but differs from it in that it also includes the concept of final causality. This leads to a realisation of different attributes of God than are accessible from the cosmological argument based on efficient causality.

I ought to note on the use of the name 'teleological argument'. This has been applied to Aquinas' fifth way, which, though different from my argument here, is still centred on the concept of final causality. However, the name 'teleological argument' has more recently been incorrectly applied to a very different argument, the argument from design. The concept of final causality was wrongly (as we have seen) removed from metaphysics in the early stages of the philosophical 'enlightenment'. Obviously, without a notion of final causality, the above argument makes little sense: unless one replaces it with something else, such as purpose. So we have 'Everything that has an efficient cause was designed by that which was its efficient cause.' That is obviously begging the question, since design implies mind, and most expressions of final causality, for example between two material objects, do not obviously depend on mind. We are left with 'Everything sufficiently complex that has an efficient cause was designed by that which was its efficient cause.' Obviously, if it is designed, then it must have a designer, and we are led to the idea of a universal mind. This argument obviously depends on what we mean by 'sufficiently complex'.

The argument from design has several flaws, the most important of which is that it accepts the enlightenment presumptions that there is only efficient causality and the universe is mechanistic. It was strongly attacked by Hume in his *Dialogues Concerning Natural Religion* and suffered another blow with Charles Darwin's *On the Origin of Species*, where it was shown, at least within the concept of biology, that there is an alternative plausible argument explaining biological complexity: Darwin's argument need not, I hasten to add, be the correct explanation of how complex life arose; the very fact that an alternative naturalistic explanation exists undermines the argument from design: it shifts the bar of 'sufficient complexity' by removing its most obvious example, biology. If there is nothing sufficiently complex in existence, then the argument of design is in tatters.

Whether or not there is indeed something sufficiently complex for the argument from design to be valid is not something I need to discuss, because, as far as any teleological argument is concerned, it is an irrelevance. Those who argue from design often implicitly accept the assumptions of mechanism: that the universe is explained by bodies in motion, efficient causality only. The laws of physics proceed independently of God, baring that He was the creator. They apply universally unless God chooses to break them. We only need to invoke God to explain those gaps in our knowledge which cannot be explained by known physical law. But as those gaps shrink as our knowledge of physical law increases, the place for God shrinks with it, until He becomes just an impotent observer, and if an impotent observer then He may as well not be there in the first place. Now this 'God of the shrinking gaps' argument fails because it presupposes that God ceases to be the explanation once the law becomes known (physical law, of course, is only a description rather than an

explanation, and a theist claims that it is a description of God's sustaining of the universe), as discussed in section 12.3.2.

The theory of evolution by natural selection thus gives a strong challenge to the argument for design. There are still gaps in our knowledge; perhaps most importantly how the first life (i.e. self-replicating molecules) emerged; but it would take a fool to say without a clear and rigorous proof that because we do not currently understand it there is no naturalistic explanation. However, evolution by natural selection is powerless against the teleological argument. It provides an alternative to design, but it cannot refute final causality because it is in itself an example of final causality; both in the broader sense, that populations tend towards meta-stable equilibrium states – and which of these states may be reached in practice is a property of the initial conditions when the environment last changed (or the final causes of the initial conditions); and in the narrower sense that it depends on individual events of efficient and final causality as genes pass from one generation to another and are mutated through unpredictable interactions with the external environment.

Now it may be argued in response that an argument from final causality is weaker than an argument from design; final causality does not necessarily involve mind, but is also evident in mindless material beings. The reader has probably already guessed my response to this objection, but for clarity it is set out below. □

CONJECTURE 13.3.2. *God possesses a will*

Proof:

(1) Final causality is manifest either in terms of the potentialities of a material object or the will of an immaterial object.
(2) Each earlier element in the teleological chain has at least one final cause.
(3) The first member of the teleological chain has at least one final cause.
(4) God is the first member of the teleological chain.
(5) God is immaterial.
(6) Therefore God has a will.

■

Discussion:
Step 1 is taken directly from the definition of will (definition 4.4.3). Step 2 says that if B_1 is a final cause of A then A has at least one final cause; this is self-evident. Therefore every element in the chain of final causality (except possibly the last) must have at least one final cause. Step 3 then follows immediately, and step 4 was shown in conjecture 13.3.1. Now it is only appropriate to say that God has a will if he both has the equivalent of final causes and is at least partially immaterial; were God material he could possess potentialities which represent the final causes in the same way as all other material objects. However, we showed above (conjecture 13.2.9) that God is wholly immaterial, and also (conjecture 13.2.3) that God lacks potentiality. Therefore, we conclude that God possesses a will.

Since God has a will, it is appropriate to say that God desires something when describing this will. Note that we have not yet shown that God is an immaterial mind; since although the will is an essential attribute of mind, it is not the only essential attribute. □

CONJECTURE 13.3.3. *Everything that exists depends on God for its existence*

Proof:

(1) A being can be either of derived existence or self-existent.
(2) If the being is of derived existence, then it can only exist if something else exists, leading to a chain of explanation.
(3) A being's existence depends on all members of the chain that proceeds it.
(4) This chain must terminate with a self-existent being.
(5) The only self-existent being is God.
(6) Therefore everything in existence ultimately depends on God for its existence.

■

Discussion:
The proof here is similar to that above for necessary beings (conjecture 13.1.10), but here we consider contingent beings as well as necessary beings. Steps 1 and 2 follow from the definitions (13.1.3 and 13.1.4), while step 3 is clear: if C can only exist because B exists, and B can only exist because A exists, then C can only exist if A (and, by extension, everything that proceeds A) exists. As argued above, the chain cannot continue indefinitely since it is instrumental, so it must terminate with a self-existent being. Since every self-existent being is necessary, and the only necessary self-existent being is God, we have thus reached the conclusion that everything only exists because God exists.

But, given that God has a will, we can go further: if God is A, then B_1 only exists because God wills it to exist, and C_{11} only exists because God wills B_1 to exist in such a way that may produce C_{11}, and so on. Therefore, beings only exist because God wills them to exist. From God's timelessness and omnipresence, we note that the way in which matter is related to God does not depend on its location within space and time; one electron relates to God in the same way as any other electron. Thus if B_1 is an electron, and God wills it to have certain final causes, then the same final causes must apply to every other electron in existence. Thus all the final causes of material objects are ultimately a reflection of God's will. If the final causes, which state how a particle can decay, are a reflection of God's will, then so is the possibility that a particle could continue in existence for a certain period of time. If the possible decays are a reflection of God's will, then the decays themselves must be willed either by God or by some instrumental agent through which God acts. Thus beings only continue to exist because and as long as God continues to will them to exist, and cease to exist as God wills them to cease to exist.

But this seems to present a problem: do we not have an alternative explanation of why beings continue to exist, and cease to exist and become something else, namely the laws of physics? If the laws of physics are thought of in materialistic or mechanistic terms, as acting independently of God, then this would indeed be a contradiction. But we need not think of them in this way; we should not presuppose mechanism if our arguments are taking us elsewhere. Rather, we should think of the laws of physics as a description of how the universe continues and evolves.

But what shall we say, then? God wills each electron into existence, and wills its annihilation. Yet God is rational, consistent and just: he treats each electron the same, no matter where or when it is in the universe, as long as all other factors (such at the matter surrounding the electron) are equal. But this has all the properties of an universal law. Indeed, we saw that the laws of physics can be reduced to an expression of the probabilities and various decay paths (or efficient and final causes) of certain kinds (or formal and material causes) of elementary particles; precisely the properties we are now attributing to the will of God. And as the attributes of the electron can be treated mathematically, we are thus led to the conclusion: □

DEFINITION 13.3.4. **Physical law**:
The Laws of physics are a mathematical description of how God sustains the matter within the universe. △

Discussion:
By 'sustains', I mean 'brings into, maintains, changes, continues in existence, and brings out of existence'. We should not, therefore, think of creation as a one-off event in the past (as we see it). God is involved in a continual act of recreation.

The consequences we can draw from this conclusion are profound, but I will discuss them in detail later (chapter 15). For now, I will just make a few quick observations:

Firstly, this position is completely contrary to mechanism. Most atheist and agnostic arguments since the enlightenment, as well as many arguments by Christians who have been caught in the enlightenment word-view, have presupposed that the operation of the laws of physics does not depend on God. But, in as much as they depend on this presupposition, these arguments can not say anything, whether in favour or against, Christianity or any other theistic religion. The difference between atheism and theism is more fundamental than the details of the laws of physics, but of what the laws of physics represent: a matter of metaphysics. It is no good arguing against (or in favour of) theism if your arguments depend on a mechanistic view of the world, since this assumption is not shared by theists. At best, you can argue against (or in favour) of deism, which also presupposes mechanism.

Secondly, to argue against theism one has to either show that the theistic conception of God cannot lead to the laws of physics as we best understand them, or that only mechanism (or some other world-view) can lead to these laws. Theism, or more precisely any one of the various different models of theism, is therefore in principle falsifiable (with the caveat that to falsify one model of theism is not enough to falsify all models of theism). From the attributes of God, we can in principle say what the laws of physics should resemble; not perhaps in every detail, but certainly in some broader aspects. If we can derive these principles, then we can compare them against what we measure, and check to see if they match. Similarly, we should apply the same tests to mechanism or any other metaphysical view which is the foundation to some form of atheism. We have, of course, already mentioned several hypotheses concerning the laws of physics that can be drawn from theism, that ought to be tested, namely that they exist, are regular, universal, unchanging in time and rational. If any of these can be shown to be false, then the version of theism I am describing here is disproved.

Thirdly, we should not be so sure that mechanism is correct. Certainly, it is a philosophy which has played an important historical role: it marked the end of the dead end of scholastic physics, and started us on the road to Newtonian Mechanics; indeed Newtonian Mechanics is a consistent mathematical expression of mechanism (as is classical relativity). For as long as Newton ruled, mechanism seemed highly plausible, if not completely certain. But we now know that Newton was wrong. This should at least make us call into question the philosophy of mechanism; certainly the naive view proposed by Galileo, Boyle and so on needs to be modified and cannot continue as it stands. As discussed in previous chapters, there is no obvious reason from the perspective of quantum physics why mechanism should be preferred and plenty of reasons why it should not .

Fourthly, a mechanist can attribute any physical process as coming about independently of God. For example, we are consistently told that evolution by natural selection describes a godless explanation for the variety of life. But this statement is based on mechanistic presuppositions. For the theist, evolution by natural selection, if it occurred, is the explanation of how God brought about the variety of life. For evolution depends on the many small details of various quantum processes, and each of those was a consequence of the will of

God. What secular evolutionists label random mutation, theists ascribe to God's sustaining of the universe. Thus, as stated above, God remains the cause of all physical processes, even after a scientific explanation is found. Indeed, it is when a scientific explanation cannot be found for some phenomena that we might need to start worrying about the existence of God (although, obviously, there are also miraculous explanations).

Fifthly, this idea that the laws of physics are a description of God's work provides a crucial difference between deism and theism. Deism sees God establishing the laws of physics when he determines the nature of each type of particle, and then He just sits back and lets the system run. For the theist, God is actively involved in the universe continually and in every place: everything we see in an experiment is an expression of God's act. I define theism as a doctrine which affirms this proposition, together with its necessary consequences. □

Objection:
There is a crucial and important objection to this discussion: final causality, at least as it is found in material objects, is a matter of probability rather than certainty. This may not hold for the immaterial analogue of final causality, the will: if one wills something, then that happens (as long as the being that wills it has the power to bring about what is willed, and is not restricted by its own limitations; this caveat does not apply to God, as we shall later see), rather than one wills A and either A_1, A_2 or A_3 occurs with various probabilities. However, for material objects, this is not the case. For example, consider the case that an electron emits a photon which decays into a pair of an up quark and anti-quark. The photon is the direct final cause of the electron; the quark and anti-quark are a indirect final cause, but may still be regarded as a final cause of the electron in an instrumental sense. However, the photon need not decay into the up quark and anti-quark pair; it may instead decay into a muon and anti-muon, for example. It is a matter of probability, not certainty.

Now suppose that God wills the initial electron into existence (let us assume for the sake of the argument that this is possible; if God created the universe out of nothing, then at some point something like this must have occurred). The electron is the final cause of God's will. The electron emits a photon, the photon decays, and it is certainly also true to say that the resultant quark and anti-quark pair is a final cause of God's initial creative act. God initially set up the laws describing the essence of the electron, and implicitly and 'simultaneously' the laws describing the essence of everything else. But this implies that God's control over nature is indeterminate rather than deterministic, which contradicts the conception of the will as determinate[11]. God does not reveal the quark and anti-quark to come into existence; he wills a set of circumstances in which either a quark or anti-quark may come into existence, or a muon and anti-muon, or something else, or nothing at all.

This seems to contradict the picture built here that everything exists because God wills it to exist. Rather, the implication is that things exist because God wills that it might exist according to a particular probability (which is determined by God). This lack of absolute control leads to something similar to the deistic God, rather than the theistic God proposed above. •

Response:
The objection misunderstands the nature of the laws of physics, at least as they are understood in the theistic world-view. The uncertainty contained within quantum theory is due to our ignorance of the mind of God. God directs the electron to emit a photon; God directs the photon to decay; but God is free to choose what the photon decays into, subject to the

[11]Note that there are two different meanings of the word determinate used in this sentence. In physics, *determinate* means impossible to predict. In philosophy of mind, *determinate* means having a fixed object or meaning.

final causes of the photon as established by God in His creation. God freely determines that it should be a quark and anti-quark; but because of our limited understanding of God's mind, and we cannot predict in advance what He would choose, the best we can make of this is to assign a probability for each type of decay. God's timelessness, omnipresence, and immutability guarantee that this probability is the same for every electron/photon in the same circumstances everywhere in the universe; and this gives the theist confidence that we can combine the various probability amplitudes for each type of particle into a universal law.

This suggests that God's involvement in the universe is not just the creation of the first particle, but God is involved in every quantum event. I have previously divided causality into substance causality and event causality. With the possible exception of the moment of creation, substance causes are material objects. The material world is sufficient to describe substance causality, and this is what physicists have uncovered with the standard model Hamiltonian. Events, on the other hand – the actualisation of potentia - are not sufficiently explained by other material beings. If the cosmos is rationally understandable, then that means that quantum events can only be explained by adding a cause in addition to the physical universe; something that is not and cannot be represented in physics. This, I suggest, is God. Certainly whatever this missing factor is, it possesses all the attributes we ascribe to God: immaterial, immutable, timeless, rational and regular, omnipresent, and so on. The same law applies everywhere. Thus God interacts with the universe not just at the moment of creation, but by actualising every quantum event.

An essential part of this response is that God has the ability to choose between one consequence and another, according to His desires; in short freedom of choice. If a theistic God had no such freedom, then the laws of physics would ultimately prove to be deterministic (i.e. every photon in identical circumstances would decay in the same way). Thus the very fact that the laws of physics are indeterminate rather than deterministic implies the following proposition: ○

CONJECTURE 13.3.5. *God has free will.*

Discussion:
Of course, this argument is only valid once we have concluded that theism rather than deism is correct. By free will, we imply that God has choices, and is not constrained in His choice of action. The deistic God need not have free will (at least, this argument does not imply it), and if this is so need not be God as usually imagined. The God of theism, however, with free will, is well on the way to being a personal being and possessing mind (indeed, mind is the only quantity we are aware of which has free will).

So this then leads to the question: theism or deism (if God has free will)/atheism (if God does not)? Are the laws of physics a description of God's acts, or of the natures of particles which continue independently of God? This is a question, I think, which can be answered by observation, as discussed in a later chapter. Certainly, one cannot say that theism is false on the basis of the laws of physics as we understand them; at least within the scope of the discussion so far. If those laws were deterministic, there would be a strong case against the theistic God; but since they are (or appear to us to be) indeterminate, this case is undermined. Secondly, theism gives us a confidence in the laws of physics that deism does not: they are unchangeable because God is unchangeable; they are rational because God is rational; we can rely on them because we can rely on God. This is immediately apparent; however, it is unclear that if the laws of physics are independent of a deistic or atheistic God that they would necessarily share His property of immutability; at least, there needs to be proof that this is so.

Thirdly, the fact of God's timelessness suggests (although I do not think that this argument is strong enough for proof) that God is theistic rather than deistic: if a deistic or atheistic God creates the first electron, then why does He not also create others at later times? What is so special about that first moment of creation? Does the existence of other matter elsewhere in the universe impede God's action in an isolated vacuum? It would be a very limited omnipotence if that were the case. Yet spontaneous creation of this form seems to deny the idea that every event can be traced back through a series of efficient causes to a single inflationary period and a single big bang[12]. I have previously argued that quantum field theory does not allow the continual creation of particles pulled from nothing; at least if this is the case, these particles do not interact with any others and cannot be said to exist.

To put it another way, we can think of each individual electron as a distortion or bump (the form) in the underlying electron field (the matter). This electron field is a single, four dimensional, object, just as the surface of a rubber sheet is a single two dimensional object. The electron field is created by God (whether a theistic, deistic or atheistic God), and the whole electron field, including its distortions, from the beginning of time to either infinity or time's end must be created instantaneously (or, rather, the equivalent to 'instantaneously' to God). The same argument applies for the fields for the other types of particles. Thus the various interactions between the particles are determined together while the indeterminate nature of physical law remains. This makes sense in theism: at each moment in the instant of creation, God chooses what a photon decays into and when (i.e. at which point on the four-dimensional fields) it decays. However, in deism (and possibly also atheism; but I chiefly discuss deism here), we require that the four dimensional fields are created once by God and then the distortions within them are allowed to evolve in time on their own probabilistically. However, this picture appears, at least at first glance, to be incompatible with God's timelessness, which suggests that everything should be created together: just as God 'observes' the matter fields as one entity across time, so He should also create them as one entity across time. If God observes the future as He creates in the past, then the evolution of the distortions in the fields must proceed towards a fixed goal. This implies (I think) that in a deistic world view the laws of physics ought to be deterministic, but experiment tells us that they are indeterminate.

That quantum physics is indeterminate implies that there is a gap in our knowledge, that goes beyond mere experimental uncertainty. In theism, this gap is because we cannot know the mind of God, which is involved in every physical interaction. In deism, the evolution of the universe is caused in part by fixed laws which act on the matter and in part by the types of matter present, and we can in principle understand both of them: once we do, there would be no gap in our knowledge beyond experimental uncertainty. There is thus no reason why a deistic physics should be indeterminate, and reasons why we might expect it to be determinate. But we observe that the universe is, at a fundamental level, indeterminate. This suggests that theism, rather than deism, is correct.

It may be objected that this argument relies on a picture of the fields as four-dimensional; rather than three dimensional fields progressing in time. However, the four dimensional picture is certainly one way of thinking about the world, and whatever metaphysics we have must be consistent with both the four-dimensional and the equivalent temporal view. I suggested in the previous paragraph that a deistic timeless God, who sits back after the moment of creation and watches the world unfold, is inconsistent with the four-dimensional view of

[12]This argument is not foolproof, and there have been suggestions that this is what happens in the universe: there is a continual inflation occurring in various corners of the universe, but which stops in other parts of the universe (such as where we are). Each individual occurrence of inflation leads to an independent bubble universe, which appears to observers within it to be similar to our universe starting from a unique big bang, but is in reality only one of many. Observers within one bubble universe cannot observe anything outside that universe, including the other bubble universes. In other words, this terminology would suggest that God is continually creating new universes within a wider multi-verse. This idea is, however, pure speculation, and unlikely to ever be demonstrated empirically.

the universe. Theism, however, is not inconsistent with the idea that beings in the universe can experience the unfolding of events in a temporal way, nor with the temporal view of physics, since just as we can consider a solid object slice by slice (without negating the bonds between the layers), particularly in the abstract, we can also consider a four dimensional world slice by slice (without negating the connections between time-slices). However, moving from a temporal world-view, where only one slice is real, to a non-deterministic four dimensional world-view is more problematic, because there is no necessary connection between the present and the various possible (but ultimately unpredictable) futures. If we conclude, as I think physics forces us to, that there is no preferred inertial frame, then it is most natural to view the fields can be considered as four-dimensional.

The appeal to what is 'most natural' (from the perspective of what physics tells us) will hopefully convince the reader that I have not proved my case. As I said, I think that the best argument in favour of theism over atheism or deism is based on observation. From here on in this chapter I will assume that the theistic view of the laws of physics is correct. However, I hope that I have shown that theism, at least to the extent of the present discussion, is compatible with an understanding of the laws of physics which is consistent with what we know of them; indeed it provides physics with a secure basis. I will propose in a subsequent chapter that if theism is correct, then the laws of physics must be similar to what we observe. I have not been able to show here that atheism is incompatible with physical law as we know of it. Theism is a possible explanation of the laws of physics; we must therefore keep it in our considerations. □

Objection:
The laws of physics are expressed without mention of God. Therefore they operate independently of God. •

Response:
Classical theism states that while God has an active role in sustaining the universe, physical events can still in part be understood through secondary physical causes. In my construction, this is seen in substance causality. So we understand the universe partially in terms of God, and partially in terms of matter. I am not advocating occasionalism, where God is the only cause. This would be a powerful objection against occasionalism, but I am not convinced that it is a strong objection to classical theism.

Physical law is constructed first of all by postulating a mapping between physical beings and an abstract representation. The mathematical objects studied by theoretical physicists are thus representations of physical beings. The objects observed by experimental physicists are also physical beings. There is no symbol in the theory representing God, but that is because of how we choose to construct the theory, and in particular the limited scope of the initial mapping. It does not necessarily imply that God is absent from reality. Similarly, the standard model of particle physics doesn't contain a graviton, a proposed particle mediating the force of gravity. That doesn't mean that there is no graviton (or other particle involved in gravity), merely that it is not incorporated into the current theory. That something is absent from the theoretical model does not mean that it is absent from reality.

But, in particular, there are further ingredients in the standard model beyond just the representations of particles. We also have the physical laws, represented by the Hamiltonian operator. Beyond that, we have quantum events such as measurements, which are not represented by anything at all in the theory, but have to be accounted for in addition to the time evolution. So these laws describing the possible interactions, and the events which convert those possibilities into actualities certainly play a role in the physical universe, but cannot be explained just in terms of the material particles. So already we have a hint that

there is something not given a specific mapping to a physical being that plays a role in the universe. There is the question of what guides the physical beings to follow physical law? What actualises the potentia, allowing things to change? And this is the role I have suggested is played by God. If this suggestion is correct, then God's interactions with the material universe are described by physical theory, but that is not immediately recognised because we use the terminology of *law* and *indeterminate quantum event* rather than specifically naming God.

The question is, given the methods and scope of science with its lack of a mathematical symbol representing God, if theism were true, how would we expect God's interactions to enter into the theory? And I would propose that they would appear precisely in terms of similar laws and models which are part of our construction. Thus describing the universe by a mathematical law acting on abstract representation of material beings is not inconsistent with God's role in sustaining the universe. It is precisely what we would expect to see if God does play the role that theism ascribes to Him. ○

DEFINITION 13.3.6. **Omnipotence:**
An omnipotent being is one that can do all things which are self-consistent and consistent with its nature.

Note that the definition here differs from a common expression, 'An omnipotent being is one that can do all things.' This definition, however, is not consistent with how Christians understand the term. This point was emphasised by the most notable theologian of the third century:

> In the next place, he [Celsus - an early critic of Christianity] objects to the statement, as if it were maintained by us, that 'God will be able to do all things,' not seeing even here how these words are meant, and what 'the all things' are which are included in it, and how it is said that God 'will be able.' But on these matters it is not necessary now to speak; for although he might with a show of reason have opposed this proposition, he has not done so. Perhaps he did not understand the arguments which might be plausibly used against it, or if he did, he saw the answers that might be returned. Now in our judgement God can do everything which it is possible for Him to do without ceasing to be God, and good, and wise. But Celsus asserts — not comprehending the meaning of the expression 'God can do all things' — 'that He will not desire to do anything wicked,' admitting that He has the power, but not the will, to commit evil. We, on the contrary, maintain that as that which by nature possesses the property of sweetening other things through its own inherent sweetness cannot produce bitterness contrary to its own peculiar nature, nor that whose nature it is to produce light through its being light can cause darkness; so neither is God able to commit wickedness, for the power of doing evil is contrary to His deity and its omnipotence[13].

I have added to Origen's definition the requirement of self-consistency, which is necessary if we are to take God and the universe as rational.

This also has biblical support, since along with statements about what God does and can do, the Bible also refers in several places to what God cannot do. For example,

> It is impossible for God to lie[14].

[13]Origen, *Against Celsus*, III.70.
[14]Hebrews 6:18.

> Let no one say when he is tempted, 'I am being tempted by God,' for God cannot be tempted with evil, and he Himself tempts no one[15].

Equally, the biblical assertion

> Beloved, let us love one another, for love is from God, and whoever loves has been born of God and knows God. Anyone who does not love does not know God, because God is love[16].

clearly indicates that God is incapable of hatred towards a substance, the opposite of *agape* love. Thus Origen's reservation, that God cannot do what is contrary to His nature, has biblical support. So, as an example, God's immutability and omnipotence are not in contradiction; God cannot change Himself, but this is not a restriction on His omnipotence when correctly defined.

This definition of omnipotence means that God cannot violate His own nature; however, for those things which are outside God, He has the ability to do what He wills with them. △

CONJECTURE 13.3.7. *God is omnipotent*

Proof:
Suppose that there is something in creation, A, (a material object) over which God has no control; which acts according to laws which are independent of God. If such a thing exists, then it would imply that God is not omnipotent. Conversely, if no such thing exists, then it would mean that everything acts according to laws which are the expression of the will of God. We will argue that, given the above, it is logically impossible for something such as A to exist. Therefore everything acts according to God's will, which is in turn only restricted by self-consistency and God's own nature. This would mean that God is omnipotent, as defined above.

(1) A, if it existed, would either still be bound by mathematical laws describing its efficient, material, formal and final causes or by no law (i.e. it has no efficient or final causes, since physical law is, in part, a precise description of every material being's efficient, material, formal and final causes).
(2) An object with no efficient and final causes does not come from anything and cannot interact with anything; therefore it does not exist.
(3) Since A is presumed to exist, it must be bound by physical (mathematical) laws.
(4) Physical law is description of God's general providence.
(5) An body under physical law is ultimately under the power of God and therefore subject to the will of God.
(6) This contradicts the premise that A is independent of God.
(7) Therefore no such being such as A can exist.
(8) Therefore God is omnipotent.

■

Discussion:
It is important to note the implicit importance of God's uniqueness (conjecture 13.2.8) in this argument. It is possible to imagine a separate chain of causality independent of God leading to a separate first efficient cause and source of the chain of final causality. But, from the above arguments, it is clear that this would be an alternative God, which is impossible.

[15] James 1:13.
[16] John 4:7

The argument itself follows directly from what was stated before. Everything that exists must be within a chain of causality, and all such chains terminate with God (conjecture 13.1.9, definition 13.1.1). We earlier concluded that God has sustaining power over everything within the chains of causality that have their root in him (conjecture 13.3.3). Sustaining power includes the ability to create, destroy, move from one potentia to another, or maintain in the same state. But since the acts of creation, annihilation, change and maintenance together describe all physical processes, it shows that God is capable of directing by His will all physical processes; that is to say all processes consistent with the nature of the various material (or immaterial) beings (with the exception of God Himself). But God Himself is the source of these natural powers, having established them according to His will at the act of creation. But since God, and the act of creation, is timeless, God cannot change or violate those powers without contradicting Himself. Thus God is capable of doing all logically possible things; with logically possible constrained by the twin chains of self-consistency and God's own nature. Since God is immutable, He cannot change His nature without entailing a contradiction; but to do something against His nature would in effect be to change His nature. For example, if God's nature is that it is impossible for Him to lie, then for God tell a lie would entail either a direct contradiction, or a change in God's nature, which is an indirect contradiction.

Therefore, with regards to created objects, God can do any logically possible thing consistent with their natural powers (which are normally immutable since they are an expression of the immutable will of God); and with regards to Himself, can do any logical thing consistent with His own immutable nature. □

Objection:
The concept of omnipotence is incoherent. For example, we can ask if God can create a weight which He cannot move. If He can create it, then His power is limited by His inability to move it. If He cannot create it, then His creative power is limited. In either case, God is not omnipotent. And, of course, this particular case is just one example of two mutually contradictory powers. Since this argument applies to any being, not just God, it shows that the concept of omnipotence is incoherent. Therefore God is not omnipotent. •

Response:
I want to begin by discussing an older example. Can God create a square circle? If yes, then we have a contradiction, since a square circle is in itself contradictory. If no, then we seemingly have a limitation of God's power. However, the answer to this dilemma has been well known for a long time. The concept of omnipotence implies not that God can create anything, but that God can create anything that is logically consistent, or perhaps that God can create anything which it is possible in principle to create. Now it is not possible in principle to create a square circle. Therefore to say that God cannot create such a thing does not contradict God's omnipotence.

Now the example of the mover seems to be different (in this example, I use 'move' in the sense of causing the object to accelerate, and 'object' in the sense of a material object, a union of form and matter). For there is nothing inherently implausible about an immovable object, at least until we stop to think about it. So the argument is presented as follows:

(1) If God is omnipotent, then He can move any object.
(2) If God is omnipotent, then He can create any object which is not self-contradictory.
(3) The idea of an object which God cannot move is not self-contradictory and thus either it can be created by God or God cannot create every non-self contradictory object.

These three ideas lead to a contradiction. The third premise either contradicts the first premise or the second premise. But whichever one it contradicts only assumes that the concept of omnipotence is coherent. Thus, if premise (3) is accepted, then the concept of omnipotence is incoherent. To defend the third point, it is observed in general that it is not self-contradictory to create an object which cannot be moved by the creator. For example, by welding together enough pieces of metal, I can create something which I cannot push or pull, even with the aid of levers or pulleys. That is true for anyone, and although we may have to use another example in the particular case of God, there seems to be nothing obvious stopping us from generalizing this principle to include God.

Except, I do not believe that the idea can apply to God. The object which God cannot move need not be merely difficult to move (as in my welded lump of metal) but impossible to move in principle, since God's strength is great enough to move anything else. Thus the question revolves around the idea of an object which is immovable even in principle. God is not the same type of being as we are; we cannot extrapolate from ourselves to God.

Now a physical object which is immovable must be of infinite mass; since everything of finite mass can in principle be moved if enough force is applied to it; indeed if even the smallest force is applied, then unless there is a counter force (such as friction), there would be acceleration[17]. However, the idea of an existent body of infinite mass is incoherent. A body of infinite mass can only arise by either combining an infinite number of particles that have finite mass, or by considering an unspecified number of particles at least one of which has an infinite rest mass. We cannot have an infinite number of particles, since that implies a countable infinity which is an incoherent notion: if we start with a finite number of particles, and keep adding to it, we will never reach infinity. On the other hand, a particle of infinite mass cannot exist: in quantum field theory, its propagator is at least inversely proportional to its mass, so any Feynman diagram involving the propagator of an infinitely massive particle must be zero. In effect, such a particle cannot interact with any other particle (since that always involves propagation), and since God cannot move it either, it cannot interact with anything, and by my earlier definition it therefore cannot exist. The idea of an infinitely massive particle (whether composite or an elementary particle) also causes problems when we consider gravity: it implies an infinite gravitational field, an infinite curvature of space time, and everything else (including space and time themselves) would be sucked into it. Since it would then be the only thing in the universe in which it subsides, it cannot interact with anything else, and once again we find that the idea of its existence is implausible. Thus the idea of an object which is immovable in principle is not coherent, and since only such an object can be heavy enough that God cannot move it, we have falsified the third premise of the argument. The contradiction in this case thus lies elsewhere, and not in the concept of omnipotence.

But surely, this is just one example? Can we say that every argument of this sort fails? I believe that there is a second response to this example which is more general. We started with the premise, *If God is omnipotent, then He can move any object.* But this is not what omnipotence implies. Omnipotence implies, *If God is omnipotent, then He can move any object which is logically possible to move.* To get from the definition of omnipotence to the first premise of the argument as stated above, we need to include an additional premise 0: *It is logically possible to move any object.* However, premise (3) implies that *It is logically possible that there is an object which cannot be moved.* The contradiction thus in practice lies between premise (3) and the hidden premise (0). Since neither of these involve the concept of omnipotence, but rather arise because the framing of the problem inherently contains a self-contradiction, we have no reason from this argument to reject omnipotence as incoherent. This objection to the argument is quite general; we can apply it not only

[17]Recall motion means change from one state to another; in Newtonian physics the states are the inertial states, and thus the Aristotelian word *motion* is equivalent to Newtonian word *acceleration* and the word has a different meaning in Aristotelian metaphysics and Newtonian physics.

to movement and creation, but to any two powers which have a mutually contradictory application. So even if an example can be found where the statement in the place of (3) above is coherent (and I am not aware of any such examples; everything I have seen leads to a self-contradiction once you stop to work through the details of what it entails, even if the contradiction is not obvious at first sight), then the argument still has no force against omnipotence. ○

CONJECTURE 13.3.8. *Only God is omnipotent.*

Proof:
Suppose that there is something which is omnipotent but not God. This would mean that it has the power to affect and alter everything apart from its own nature, which would include God. But this is impossible, since this would imply that there was something within God which would be caused by something outside God, which violates either His simplicity or His self-existent nature. Therefore only God can be omnipotent. ■

CONJECTURE 13.3.9. *If there is an omnipotent being, that being is God.*

Proof:

(1) An omnipotent being is capable of doing all things that are logically possible.
(2) Every omnipotent being must sit in a chain of efficient causality.
(3) The omnipotent being would either be the terminus of the chain of efficient causes, or itself has an efficient cause.
(4) If it was the terminus of the chain, then it would be God.
(5) If it was caused by something else, then it is logically possible that it can be caused.
(6) No being can be its own efficient cause.
(7) An omnipotent being cannot be a secondary link in the chain.
(8) Therefore, if there is an omnipotent being, that being is God.

■

Discussion:
Point 1 is a restatement of the definition of omnipotence (13.3.6). Point 2 follows from the principle of causality (4.5.14). Point 4 follows from the definition (13.1.12) and uniqueness (13.2.8) of God. Points 5-7 refer to the second option from point 3, and demonstrate that an omnipotent being cannot be something other than the terminus of the series of causes. Since it can't cause itself, and it isn't something that by its nature cannot have a cause, it could not be omnipotent. This contradicts the original assumption that the being is omnipotent, thus showing that the being cannot be a secondary member of the series. We are left with the conclusion, that either no omnipotent being exists, or that any omnipotent being must be God.

We could have defined God as the omnipotent being, and this definition would thus be equivalent to definition 13.1.12. Thus God's omnipotence and his necessary self-existence are two aspects of the same thing; we distinguish them only on account of linguistic convenience: when discussing God's power, it is more convenient to talk in terms of omnipotence because the word directly relates to power; when discussing His immutability it is more convenient to discuss of His necessity and self-existence. But in reality, there is no difference between the two. God's necessary self-existence entails His omnipotence; God's omnipotence entails

His necessary self-existence. □

CONJECTURE 13.3.10. *The form of a contingent substance is implicitly present within its efficient cause.*

The definition of 'implicitly present' will be described in the proof and discussion below.

Proof:

(1) Every existing contingent substance is the final cause of another substance.
(2) The final cause of a being thus indicates both the form and matter of the effect.
(3) The potentialities of a being include form.
(4) The form of something that is caused is contained within the essence of that which causes it.

■

Discussion:
The first point is directly taken from the definitions of contingency (13.1.5) and final cause (4.5.10). The second is also clear; if a match ignites when it is struck, then part of the definition of a match is its final cause: that it creates fire when rubbed against a rough surface. This only makes sense if we have first defined what fire is. In other words, the definition of the match, to be complete, must also contain the definition of 'fire'. Thus the essence of the match includes the description of the form of fire. Fire is not within the match actually, but is within the match either potentially or virtually. The same is true for any other pair of cause and effect within the material world. Every complete description of something, the essence, includes a description of its final causes; this cannot be the substance (form and matter) of the final causes, otherwise the matter of the efficient cause would take on two different forms, which is impossible (the same piece of matter would be two different things at once). Therefore, essence includes a description of the form of its final causes; the form of the final cause is contained within the efficient cause. For a material efficient cause, this is either as a potentia or virtually (definitions 4.5.2 and 4.5.3); for an immaterial efficient cause (should there be such a thing) it is either as a potentia or part of an intellect; since the intellect is the a way an immaterial being can grasp an abstract concept such as form. We saw earlier that form can only be exhibited actually (which is not the case here), potentially or within an intellect. I will define 'implicitly present' as meaning as either as present either within the potentia of an object or grasped within an intellect. □

CONJECTURE 13.3.11. *The form of all things is implicitly present within God.*

Proof:

(1) The form of a contingent substance is implicitly present within its efficient cause.
(2) It is also present within the efficient cause of its efficient cause, and all its causes back to the first cause.
(3) The first cause is God.
(4) Therefore the form of all things is implicitly present within God.

■

Discussion:
The first step of this argument is the conclusion of conjecture 13.3.10. Point 2 follows by the same reasoning, as if B is a final cause of A and C a final cause of B then C is indirectly a final cause of A. To fully define A, one must first define B, and from there one must define C. Thus C exists implicitly within A, and from A we trivially can extend this argument back to the first cause, God. Since this argument follows for any material substance, the form of every material substance must be implicitly present within God. □

CONJECTURE 13.3.12. *God has an intellect*

Proof:

(1) The form of all things is implicitly present within God.
(2) Form must be present either actually, potentially, virtually or intellectually.
(3) The form of a material object is not actually present within God.
(4) The form of a material object is not potentially present within God.
(5) The form of a material object is not virtually present within God.
(6) The form of a material object must be intellectually present within God.
(7) God possess an intellect.

■

Discussion:
Point 1 is taken from the conjecture 13.3.11; while 2 is taken from the definition of intellect (definition 4.4.2, see also 4.5.2, 4.5.3 and 4.5.4). It is clear that the form of, for example, a teapot cannot exist actually within God, otherwise God would be a teapot, which is obvious nonsense. Neither can it exist potentially, since God is pure actuality. Nor can it exist virtually, since for God to create an electron would then involve a change in God, which is impossible (conjecture 13.2.2). Therefore it must exist intellectually. Since only an intellect can grasp objects intellectually, it follows that God possess an intellect. □

DEFINITION 13.3.13. **Knowledge**:
To know an abstract concept is to possess it indubitably, precisely and accurately within one's intellect.

Indubitably means that one does not reasonably doubt the information; otherwise it is not knowledge but belief or opinion. Precisely means that the knowledge is not constrained within unreasonably large uncertainties; for example, the statement $2+2$ *lies between* 0 *and* 25.7 does not count as knowledge according to this definition because, although accurate and held without doubt, if it is all that is understood about $2+2$ then it it unreasonably imprecise. Accurately means that it is in accordance with reality. The belief that *All cats have six legs* is precise and may be held indubitably, but is false knowledge because it is inaccurate.

Obviously the use of the word *unreasonably* means that this set of definitions is not precise. Ideally, knowledge should be perfectly precise; given, however, that this is impossible for us, the concept needs to be loosened. *Reasonable* means in view of the amount of information which is in principle available to the intellect; what he could have found out if he had submitted all the resources available to him to seeking it. I will label this sense of the word, which deals with abstract statements and their correspondence to reality, as *propositional knowledge*, and treat it as the primary sense of the word.

A further sense of the word know is *experiential knowledge*, for example to know what it feels like to sit down. This is different from knowing the mechanics of how a body moves when it sits down, the various reactions of the nervous systems and so on, all of which can be reduced to propositional knowledge. We are still left with the feeling or act of experiencing itself. I am unsure about whether this can be reduced to propositional knowledge or not, and will leave this as an open question; to answer it would be to delve deeper into neuroscience and the philosophy of mind than I am capable of doing.

There is also *participatory knowledge*, which is the knowledge we gain by being something. I know what it is to be myself, but not through reason, nor through experiences, but simply through being me. Obviously *participatory knowledge* is wholly subjective: only I can know what it is to be me through participation; only a woman can understand what it is to be a woman through participation. A man may think that they know what it is to be feminine, since our irrational minds can all think absurdities, but they cannot truly know what it is to be female given that they are not one.

There is a different (though related) sense of the word know which means to know a person; i.e. to have met and observed some of that person's characteristics and character. But this goes deeper than that; when I might say 'I know Peter' I mean more than just 'I know certain facts about Peter'. This I will refer to as *acquaintial knowledge*. It is important to recognize that the word 'know' in this sense means something different from what is meant when we discuss propositional knowledge. Much as I hate to admit it, the French knew what they were doing when they distinguished between *connaitre* and *sais*.

In my discussion of knowledge, I concentrate on the first of these types of knowledge, which I term *propositional knowledge*; my definition of omniscience focuses on this type of knowledge. I will briefly discuss experiential and participatory knowledge in the context of various objections; I will not consider whether or not God can have acquaintial knowledge, and will not include this as part of my definition of omniscience. That does not mean, of course, that God cannot posses this type of knowledge (as a Christian, I believe that He does), only that I do not discuss it in this work. Δ

CONJECTURE 13.3.14. *God knows all things which it is logically possible to know and which are consistent with His nature.*

Proof:

(1) All knowledge is either of those abstract ideas unrelated to forms, or the forms of material objects.
(2) To understand an object is to grasp its form within the intellect.
(3) To fully know about an object is to understand its form perfectly.
(4) To understand all forms means to understand all relevant abstract ideas; since the abstractions are made manifest in one form or another.
(5) God's knowledge is not bound by time, space, or any other constraints except self-consistency.
(6) The form of all things is intellectually present within God.
(7) This form is manifested perfectly.
(8) Therefore God knows all things related to the forms of actual objects.
(9) Therefore God possesses all knowledge relevant to this universe.
(10) God need not have created this universe.
(11) God possesses this knowledge regardless of whether He created this universe or not.
(12) God possesses all knowledge relevant to any logically possible universe.

(13) God knows all things which it is logically possible to know.

■

Discussion:
Point 1 Follows from the definition of knowledge (13.3.13), and that every abstract idea must either be the form of a material substance, the form of an immaterial substance, or something else. Point 2 is from the definition of intellectual understanding (4.4.2). Point 3 follows from specialising the definition of knowledge to the particular case of forms.

Point 4 is a little more subtle. The claim is that abstract ideas concerning, for example, triangles can in principle be derivable from the forms of triangular objects. For example, one can verify the mathematical relations concerning triangles by knowledge of all possible triangles; that is all possible potentia of a triangular object. To understand equilateral triangles, one must understand angles; to understand triangles of a particular size one must have prior knowledge of length; from these one gains an understanding of the abstract concepts of trigonometry, and subsequently geometry. Of course, we do not prove the properties of triangles from understanding all possible triangles; what I am saying is that knowledge of those properties is either a prerequisite or corollary of any understanding of the form of a triangular object. To understand objects with extension one must understand extension, and from this an understanding of the abstract concept of number. To fully understand electronic devices one must understand logic gates; and to understand the purposes of logic gates we must previously know the abstract concepts of binary arithmetic and the basics of (modern) logic. To fully understand men, one must understand the abstract concept of language; for an essential part of the nature of men is that we communicate, and language is the means that we do so. To fully understand all possible forms existent in a given universe, then requires a great deal of understanding of abstract ideas which are not forms. I have not shown that all abstract ideas need to be known for there to be a complete knowledge of form; but certainly all abstract ideas which are manifest within the material universe need to be known, and this is what I mean by all *relevant* abstract ideas. In other words, knowledge of forms presupposes knowledge of the various mathematical and logical abstractions, so if one perfectly knows a form, then one would also understand those abstractions as well.

Point 5 follows from God's necessity, which was part of the definition of God above (definitions 13.1.6 and 13.1.12). If God's knowledge is restricted to one time (for example), then God must be restricted by something external to him or there must a a boundary or sense of time within God. But, as shown above, this is impossible. This point is important because otherwise we could say that God knows all things in the present but not the future, at one place but not other places.

Point 6 was shown in conjecture 13.3.11. By *perfectly* in point 7, I mean precisely and accurately, and this is clear because existent objects have a precise form, and since this form in actuality can be traced back to a potentia in some other object and ultimately to God's intellect the form in God's intellect would have to be precise, otherwise the potentia and then the actual form would not also be precise. Points 8 and 9 follow directly from the previous statements.

We now come to point 10. This follows from the freedom of God, established above (conjecture 13.3.5). It is not essential to God's nature that He should have created this universe; His existence in no way depends on the universe because if it did God would not be self-existent, which is contradictory to the definition of God (definitions 13.1.4 and 13.1.12). Thus God need not have created this universe. But, since God is immutable (conjecture 13.2.2), and cannot be not changed by anything outside Himself, His knowledge of the forms and abstract concepts of this universe cannot be affected by whether or not God actually did create this universe. If it did, it would once again imply that some aspect of God was conditional upon something outside of God. The only knowledge God gains from the creation of the universe is that some substances are actual rather than potential;

but since God is the cause of their actuality, this knowledge does not come from outside God but from within God and thus does not violate His self-existence.

Now, what is true for our universe must also be true for any other logically possible universe[18] which God could have created; regardless of whether God has created them in practice[19]. Given that every logically possible form must occur in at least one possible universe, and every possible applicable non-form abstract concept[20] must be known prior to the knowledge of at least some logically possible form, we may conclude that God knows everything that it is logically possible to know. □

Comment:
Note that this argument only applies to propositional knowledge. It also only applies to statements which are both true and necessary. Statements can either be true, false, or neither true nor false. For example, the statement "2+2=4" is true (given the standard definitions), "2+3=19.5" is false, and "It is now time for dinner" by itself is neither true nor false, without additional information (such as when 'now' is and when dinner will be served). The proposition "It is now time for dinner" is only true or false in combination with other statements, so it is only contingently true or false. Such contingent truths are only contained within omniscience as far as they relate to logical chains of reasoning. For example, if the statement 'If A is true, then B is true' is logically sound, then this propositional structure is known as part of omniscience, but the statement B is true absolutely (or B is not true absolutely) is not part of omniscience because B is only a contingent truth. □

DEFINITION 13.3.15. **Omniscience**:
To be omniscient means that one perfectly knows all things, subject only to the constraints of logical possibility and self-consistency. △

Discussion:
From the previous proof and this definition, it is clear that God is omniscient. In this definition, *Know* is in the sense of propositional knowledge, and excludes contingent propositions except as discussed in the above comment. □

CONJECTURE 13.3.16. *Only God can be omniscient.*

Proof:
Suppose that A is the cause of B and also may be the cause of C, while there are aspects of the form of C not included in B either potentially or intellectually.

(1) A possesses the form of B either potentially, virtually or intellectually.
(2) A may also contain other abstract ideas either potentially. virtually or intellectually.
(3) If B cannot cause A, then the form of A does not reside in B potentially.

[18]I mean by this every possible universe which God could in principle create; this is different from the sense of 'all possible worlds' used by contemporary philosophers, which includes the 'world' where nothing exists, or there is no causation through conservation of energy and momentum, and therefore God also logically need not exist.

[19]We do not and, I think, cannot know if our universe is the only actual universe; but that is irrelevant to the point in hand.

[20]I am excluding, for example, concepts which are self-contradictory or cannot be applied in any way to any possible material object. I am including within God's knowledge any self-consistent concept residing or which might reside in the thoughts of mathematicians and philosophers.

(4) If A is also the cause of C, then it must contain abstract ideas that not contained within B potentially.
(5) Therefore, the form of A cannot fully reside in the actuality or potentia of B.
(6) Intellectual knowledge originates either in creation, participation, sensation or reasoning.
(7) Something other than God cannot gain full knowledge of God through creation.
(8) Something which is not God cannot contain God's form by participation.
(9) One can only sense material objects.
(10) Something other than God cannot gain full knowledge of God through sensation.
(11) One cannot, by reasoning, deduce an infinite amount of information.
(12) To understand God fully requires an infinite amount of information.
(13) Therefore, something other than God cannot gain full knowledge of God through reasoning.
(14) Therefore nothing outside God can fully possess God's form within its potentia, actual form or intellect.
(15) Therefore only God can be omniscient.

■

Discussion:
This proof is in two steps. The final goal is to demonstrate step 14, which I divide into two separate concepts; firstly something which possess the form of God either actually or potentially; secondly whether something can possess God's form intellectually. For example, in the first case, if God's form could be summed up by knowledge of XYZ, and there exists something which is actually X and could potentially be Y and Z, then that something would possess all of God's form. This does not, however occur. For example, we may consider that in some other possible universe there is some particle which is identical to the electron except that it's rest mass is the same as some particular elephant (in our universe). God certainly has knowledge of such a particle, since it is logically possible, but it does not exist actually or potentially in this universe because such a particle does not occur in this universe. Therefore, the entirety of God cannot exist actually or potentially within any being other than God. We could also argue that if God existed potentially in another being, then it would in principle be possible for that being to change into God, which is clearly false given that God is necessarily immutable. This leads us to step 5.

We now come to intellectual knowledge, which perhaps ought to be our main focus. In step 6, I discuss creation, participation, sensation and reasoning. I exclude imagination, because what is imagined has no actual existence (unless it is created, or the imagination is inspired by participation, sensation or reasoning). Creation is the process by which God understands form: He imagines it, and then brings it into existence. However, nothing can know God in this way because God is not created (point 7, see also conjecture 13.1.9). Participation refers to that sort of understanding we gain by being something; however, since only God is God, only God can gain knowledge of Himself in this way (and this, of course, is how God avoids this proof and can understand Himself). This is point 8.

Next we come to sensation. This is the acquisition of knowledge through our senses. However, we can only directly sense that which touches us (for example, a photon of light hitting our eye), we means that we can only sense material objects, or interactions between the immaterial and immaterial world. However, there are (an infinite number of) aspects of God which are fully immaterial. Therefore, we cannot gain full knowledge of God through sensation.

Finally, abstract reasoning builds on premises, and increases knowledge one step at a time. However, one can never reach infinity one step at a time. However, full knowledge of God requires knowing all that God knows, including every logical possibility, and every aspect of this universe: an infinite amount of information. Therefore, one cannot deduce

everything about God through reason. Thus nothing outside God can fully know God, either potentially, intellectually or by participation or a combination of them. But to be omniscient implies full knowledge of God in at least one of these ways. Therefore, only God is omniscient.

Note that this does not mean that we cannot know accurately about God in part, but that our knowledge of God must be incomplete or imprecise. If we could not know anything about God, then this whole exercise would be self-contradictory. However, even though what is uncovered in this chapter is (I believe) accurate, what I have stated is not perfectly precise and it is certainly not complete. Thus to say that we can know some things about God with reasonable accuracy does not contradict the statement that we cannot know everything about God with perfect precision.

This means that to be omniscient is logically equivalent to the definition of God, definition 13.1.12. God is not the necessary self-existent and the omniscient one (which violates God's simplicity), but they are simply two different ways of saying the same thing. Equally, omniscience is equivalent to omnipotence and omnipresence. □

Objection:

(1) There are certain things which it is only possible to know by experience.
(2) There are certain experiences which an omnipotent being cannot have.
(3) Therefore an omnipotent being cannot know everything.
(4) Therefore an omnipotent being cannot be omniscient.
(5) God is omnipotent, as shown in conjecture 13.3.7.
(6) Therefore God cannot be omniscient.

Examples of things only knowable by experience which an omnipotent being cannot experience include what it feels like to experience fear, or what it feels like to not be an omnipotent being, and so on. This is based not on an understanding of fear as an abstract concept, but the direct sensation. Clearly, this is not understandable except as it is sensed. Fear is inconsistent with omnipotence because fear is a reaction to danger, and nothing is dangerous to an omnipotent being. There are other variants of this argument: for example, one cannot be a perfectly moral being and still understand what it feels like to experience lust, sloth or envy. Therefore one cannot know this, and therefore one cannot be both a perfectly moral being and know everything. •

Response:

There are several possible answers to this. The first is to note that in Christianity, God came down as man, and shared the experiences and temptations of man. As such, He experienced many of the feelings directly. Does this mean that God logically had to become incarnate to complete his knowledge, and is this inconsistent with the notion that the incarnation was to bring salvation to mankind? I do not think it is inconsistent, because it is not inconceivable that there could be numerous reasons for the incarnation. However, this answer is clearly inadequate; firstly because since Christ was without sin[21], this does not fully answer the moral version of the argument; secondly, let us consider another race of beings, the Quingoids from the planet Qog, who, if they existed (which they don't: they are logically possible -- nothing inconsistent in their nature, but happen not to be on the planet Qog, partly because there is no planet Qog in the first place) would experience the feeling of Quinging, which is unique to their (non-existent) race. Now, if the incarnation is the only response to this objection is valid, then knowledge of the experience of Quinging would only be possible for Quingoids, so to experience this God would have to also become incarnate as a Quingoid.

[21]Hebrews 4:15.

But this is impossible, since there are no Quingoids. Thus God's knowledge would have to exclude what it feels like to be Quinging, and would thus be incomplete. In this way, the incarnation is only a partial but not complete answer to the objection.

We may also challenge whether or not it is necessary to be in personal danger to experience fear. For example, when I watch a horror film, I might experience fear through my empathy with the characters depicted on the screen. Perhaps God could experience fear in the same way? Even if this is right, it is not clear that it answers all the possible forms of the objection. Also, it not clear that God would be able to experience fear in this way: I can experience fear when watching the film because I have been in danger myself, and thus know what sort of experience the characters are going through.

Another possible response is to say that God's knowledge is limited to only propositional knowledge – after all, this is all that is proved above. If experiences can be reduced to propositions, then ultimately they can be known as abstract concepts, and so the premise in the objection that there are things which it is only possible to know by experiencing them is false. If experiences cannot be reduced to propositions, then experiential knowledge is not part of God's omniscience as defined above, and thus there is no contradiction. However, this response also runs into difficulties with the incarnation; but this time the opposite difficulty. For God as man certainly felt some things, and thus possessed some degree of experiential knowledge, if not complete experiential knowledge. Thus, if the incarnation is correct, then it must be possible for God to understand at least some forms of experiential knowledge. Therefore, God's knowledge is not just by logical necessity limited to propositions; which means that the obvious solution to God's lack of certain types of experiential knowledge (that it is logically impossible for God to know that) is invalid.

So let me tackle this objection in a different way. Suppose there is a man – let's call him David – who has experienced the feeling of fear (and again it is the feeling of fear and the experience, not an abstract understanding of what fear is that is at stake here). Perhaps he has to present an untenable proposition before those most clear minded of people, the society of Thomists, and is worried about what little will be left of his academic reputation when they are done with him. In any case, he experiences something. Now what he experiences is not the feeling of fear in the abstract, but the feeling of fear as it is experienced by David. In all this type of knowledge of internal experiences, we cannot separate the experience from the one who is experiencing it. Now, God knows David fully. For example, if David's self-awareness just comes from the electrical signals in his brain, then God will know all about those, and how they lead to David's self-awareness, and will thus have a complete knowledge of the self-awareness. If the self-awareness arises from electrical signals in the brain plus something else, then whatever that something else is it is also ultimately caused by and sustained by God, and we still reach the same place. Thus God knows fully the causes of David's consciousness, and through that the manifestation of David's consciousness and through that how David experiences his own consciousness. God knows David's understanding of himself fully. Now, if God knows David fully, then he also knows how David experiences fear. Thus while God cannot gain direct experiential knowledge of certain quantities, he can know it indirectly through complete knowledge of the feelings and experiences of others.

Now, it is also clear that God's knowledge of the experience of fear is greater than David's. For David only knows the experience of fear as he himself has felt it. There is no obvious reason to suppose that someone else, say Saul, experiences fear in the same way that David experiences fear. Nor is it it clear that the experiences of some non-human animal are the same as either David or Saul's. David's experiential knowledge of fear is thus limited to his own experiences; Saul's to his own experiences; an antelope's to its own experiences. Yet God's experiential knowledge of fear encompasses all of them. Nor is it necessary that a being exists for God to understand how it would experience fear if it existed: for God knows the complete form of the being, and is able to fully deduce from that its experience

of self-awareness. And this brings us back to the Quingoids. For though they don't exist, God still understands their form, and can deduce from that what it would feel like for any one of them to experience Quinging in the same way that He can deduce how David feels when he experiences fear from knowledge of David's form. Thus God's knowledge of these types of feelings is not limited to those feelings that have been or will be felt in practice, but encompasses everything that is possible in principle to be felt.

Therefore the objection fails, and God's omniscience is not limited by His other properties in this way. Indeed, we have just shown that God's omniscience is not only limited to *propositional knowledge*, but also all logically possible *experiential knowledge.* ○

Objection:
God's omniscience is inconsistent with His timelessness. For an important part of knowledge is the concept of *now*: that an event is happening now, that we are going to the supermarket now. Similarly, we also require tensed language: it will rain tomorrow; it did rain yesterday. These statements are either true or false; yet for a timeless being they are meaningless. Therefore there is some knowledge which a timeless being cannot understand. Therefore God is either not omniscient or not timeless. •

Response:
We ought to separate two separate senses within a statement that in some way refers to the present time. Firstly, it contains propositional content. For example, the statement 'It is now raining,' contains information which is either true or false. Secondly, there is the experience of *nowness*, one particular moment being the present, and other times either being past or future in respect to it.

It is important to observe that the phrase 'It is now raining' is meaningless without additional information. For example, it implicitly contains information which relates to the speaker, namely where and when he is when he utters the statement. 'It is now raining' means 'It is raining simultaneously with these words or thoughts in the same place in which they are being spoken or thought.' The word *now*, as far as its information content is concerned, is no more than a shorthand for all of that longer sentence baring the reference to rain. There are two events, the words being spoken, and the rain, and the word *now* just describes the time that elapses between them. In this way, every statement which contains reference to the present time can be re-expressed in entirely equivalent (as far is its information content is concerned) tense-less or timeless language. We see this from physics: there are two constructions of physics, the four dimensional approach where we treat all of time together, and the temporal approach where we treat each time slice separately, and they are entirely equivalent: every physical observable considered from a temporal representation has a one to one correspondence with an expression of the observable from a four dimensional perspective. It is just a matter of using different words to convey the same meaning: the words are unimportant – just an arbitrary convention, the meaning is unchanged.

That leaves the experience that one moment of time is singled out as the present. Now, it is true that a timeless being cannot have direct experience of this sort of understanding; but nonetheless, it can understand it indirectly. I understand what it is to be a temporal being, to experience one moment of time following another; and God understands me fully. Therefore God also understands fully my experience of the passing of time, and thus this knowledge is accessible to God even without a direct experience of it. God's omniscience includes how I experience life, and therefore does not lack knowledge of the experience of temporality, even though He Himself is timeless.

Therefore this is no reason to consider timelessness inconsistent with omniscience. ○

Objection:
I have the knowledge that *I am sitting on a chair.* Nobody else can have precisely this knowledge – the best they can say is that *He is sitting on a chair.* However, this is different; only I can have first person knowledge about myself. Clearly my friend can also know that *I am lying on a bed,* and only he can know this in a first person sense. Since there is knowledge which only I can know and knowledge which only my friend can know, neither of us can be omniscient. By extension to all beings, we see that omniscience is an incoherent concept. Therefore God cannot be omniscient. •

Response:
This objection is similar to the previous one, and the response is also similar. There are three aspects of the knowledge contained within the statement *I am sitting on a chair.* The first is propositional knowledge; the knowledge that at a certain time a certain individual is sitting on a certain chair. But this is in principle knowable to everyone. The second part is the experience of sitting on that particular chair at that particular time, which is known only to me and to anything that knows all the impulses in my brain perfectly. Neither of these are contrary to God's omniscience, for reasons already given. If there is something left after these, then it is participatory knowledge – knowledge which depends on the knower. However, this also is not an issue with God's omniscience. For omniscience implies only knowing those things which it is logically possible to know given one's nature, and it is not logically possible for God to truthfully know 'I am Peter' since He is not Peter. Thus any purely participatory knowledge which cannot be reduced to propositional or experiential knowledge is not contrary to omniscience since it lies outside the definition of omniscience. ○

CONJECTURE 13.3.17. *God is essentially an immaterial mind.*

Proof:

(1) A mind is that which possesses intellect, perception, will and freedom.
(2) God possesses intellect essentially.
(3) God possesses will essentially.
(4) God is able to perceive other beings.
(5) God possesses freedom.
(6) God is essentially immaterial.
(7) God is an immaterial mind.

■

Discussion:
Point 1 is the definition of mind provided earlier (4.4.1). Point 2 was partially shown above (conjecture 13.3.12); the remainder is that intellect is essential to God's nature. However, since God's omniscience is equivalent to His essence (conjecture 13.3.14) His possession of intellect must be essential to Him. God's free will (points 3 and 5) was also shown above (conjectures 13.3.2 and 13.3.5). Similarly, from the discussion of omnipotence (13.3.7), we see that the free will is essential to God's nature. God's ability to perceive other beings is not through senses, as we do, but through the act of continually sustaining them (conjecture

13.3.3). As already seen, God must be immaterial (conjecture 13.2.9); and therefore we are left with the conclusion that God is an immaterial mind.

Now it may be objected that the mind of God is very different from our minds; for example, God is timeless, omnipresent, omniscient, and so on, while we most definitely are not. But this is not a blow to Christianity; the Christian states that while the mind of God resembles ours in certain respects[22] it also differs from us in important ways[23]. We are similar in the respects that we are capable of understanding abstract concepts, have some freedom of choice, and can use our knowledge and freedom to direct our actions. However, that does not mean that these abilities are manifested within ourselves and God in the same way. The same word *mind* is used to describe ourselves and God analogously rather than identically. God's mind is the primary use of the word; we may use the same word to describe some of the same abilities as God, even though there are large differences between how God thinks and how we think (and even between what thinking means within God and within us). □

DEFINITION 13.3.18. **Create:**
A creator is a rational agent who brings something else into being, either by rearranging other matter, or from nothing (*ex nihilo*). The act of creation, which is what is carried out by a creator as a creator, is entailed in the verb *to create*. △

CONJECTURE 13.3.19. *God is the creator of everything except Himself.*

Proof:

(1) A creator is a rational agent whose will is a primary efficient cause of the object created.
(2) God is a rational agent.
(3) God is the primary efficient cause of everything except himself.
(4) God is the creator *ex nihilo* of all matter in the universe.
(5) God, in his act of sustaining the universe, can reorder matter as He wills.
(6) God created the universe out of nothing, and has the ability to create by rearranging matter.
(7) God is the creator of everything.

■

Discussion:
The first point is from the definition of the creator (13.3.18), while the second was established in conjecture 13.3.17 and the definition that anything that has a mind is a rational agent. From the cosmological argument (conjectures 13.1.9 and 13.1.10), we see that God is the ultimate cause of everything, including time, space and all matter (as matter cannot

[22]See, for example, Genesis 1:26: *Then God said, 'Let us make man in our image, after our likeness. And let them have dominion over the fish of the sea and over the birds of the heavens and over the livestock and over all the earth and over every creeping thing that creeps on the earth.' So God created man in his own image, in the image of God he created him; male and female he created them.* The image and likeness states that there is some resemblance between God and ourselves. This resemblance cannot be physical, since God is immaterial, but must refer principally to our immaterial aspects, our minds.

[23]See, for example, Isaiah 55:8: *For my thoughts are not your thoughts, neither are your ways my ways, declares the Lord. For as the heavens are higher than the earth, so are my ways higher than your ways and my thoughts than your thoughts.*

exist outside time and space). Therefore God is the creator of the universe *ex nihilo* (either directly or through an instrument). Equally, as God upholds the universe, and can bring things into existence, continue them in existence, change them, and remove them from existence at will (conjecture 13.3.3), he can also create by rearranging matter. Therefore, both as the first cause and by direct intervention, God can be regarded as the creator. Since all things are members of chains of causality which ultimately lead back to God, we can say that God is the creator of everything. □

13.4. The Persons of God

DEFINITION 13.4.1. **Hypostasis**:
A hypostasis is an individual inseparable building block of a substance. △

Discussion:
For a complex substance, which one can separate into parts, each of those parts would be a distinct element of it. However, I will use the English words 'part' to denote these differences within the substance. But what of a simple substance? I defined a simple substance as one not composed of separable parts. But this does not exclude that it has distinctions within it; only that these distinctions are not separable, in that if one is there the other (logically) necessarily must be present; and if the other is there the one must (logically) necessarily be present. We can think, perhaps of the two ends of a finite string (although this analogy is flawed, as strings that we see are not simple substances; I consider a string as a purely one dimensional object, so it cannot cross itself or join itself in the middle). A string can be circular, and have no ends, or it can have two ends, but logically a string cannot have a single end. Even if the string is cut, each end of the string will still be connected to another end. Thus, ontologically, the two ends of the string cannot be separated from each other (they are separated by distance, but that will not be possible within God). Both ends of the string are part of the same substance; logically one cannot have one without the other. However, they remain distinct from each other. To describe this sort of phenomena, we need a word which indicates a distinction within a substance which is not caused by a mixture of separable substances; and for this I am using the Greek word *hypostasis*[24]. I shall use the word *element* to describe either a *part* or a *hypostasis*. □

DEFINITION 13.4.2. **Person**:
A person is an individual element of a rational nature. △

Discussion:
This definition is different from how we use the word in modern English; however, it is the concept used by the Latin theologians, and the English word 'person' has traditionally been used to translate that concept. I will therefore use the word in this traditional rather than modern sense, as the modern sense does not accurately convey the idea that I need.

So what does this definition mean? I will start with *of a rational nature*. This implies a substance that essentially contains mind. For example, humanity as a species is a substance

[24]Note also my use of difference to indicate a mixture of substance and distinction to indicate hypostasis within a substance. Also, the picture of the string has some similarities with the Christian doctrine of the Trinity, but also, as we shall see, differs from it in an important respect.

of a rational nature[25], since the tendency to reason is essential to the nature of men[26]. Now, individual men inherit the property of rationality from the definition of their species. A man is rational, not because he is rational in himself, but because he is a member of the species man. However, the rational nature of the species man is not derived from the higher substance, since the next level up from man is 'animal' or perhaps 'primate', but it is not essential to the nature of primates to be rational agents. Thus man as a species is the most general level in this tree where the concept of rationality is part of its essence. Individual men are individual elements of this nature[27], and it is therefore to individual men that the word 'person' most properly applies: not to groups of men, nor components within a man.

So, for example, a rock is not a person because although it is an individual element it is not of a rational nature. A football team is not a person because although composed of a rational nature (at least, in most cases – sometimes England's performances leave me in doubt) it is not an individual. A man's heart is not a person because it is a part of a rational substance and not a rational substance in itself. A individual man is, however, an individual element of a rational nature, and thus a person.

The word person is thus used as much as a mark of distinction as to describe the nature of something.

Now, what if there are distinctions within God? What follows from the nature of God – eternity, omnipresence, majesty, being the creator, uncreated, not fully understandable for other beings, omnipotence, and other any attributes which I have not discussed here, would be equally shared between each of the hypostasis. God is of a rational nature, and does not derive that nature from some more general substance (since God's nature is not derived from anything else). Therefore, we can rightly describe these individual hypostasis using the word person, as defined above. Thus, while God is simple in substance and attributes, this does not preclude that there are distinctions within God; united in God's nature, but distinct in personhood.

The only question is whether there are indeed such distinctions within God, and, if so, what distinguishes the individual hypostasis. This cannot be a matter of different substance or different parts, but of some necessary and inseparable connection. □

DEFINITION 13.4.3. **Relation:**
Relation indicates a necessary connection between two different or distinct elements. △

---

[25]To those who object by saying that certain individuals may not possess rationality; either due to disease, damage, drunkenness, old age, young age, or by being part of the social science faculty; I will respond that this is a misunderstanding of what is meant by nature or the substance of a species. The rational nature indicates what men naturally become if there are no obstacles preventing it. For example, a human fetus with limited thoughts is an early stage of development of an adult human with (usually) more sophisticated thoughts; they are, however, of the same substance, both human; indeed possibly both the same individual only distinguished in time. One thing that differentiates man from monkeys is our rational nature; it is part of the essence of man.

Therefore all men are of a rational nature, even though some, such as fetuses or those in a coma are not yet able to take advantage of it at that time (but will at other stages of their life, if they are allowed to continue in their natural development), and others, such as certain politicians, seem to have been trained not to use it.

[26]The word 'man' and 'men' traditionally has two senses in English; firstly as a specific reference to male representatives of humanity; secondly as a general term which refers to humanity as a whole, both male and female. English, unfortunately, does not have the distinction between *aner* and *anthropos* present in Greek, but the same word 'man' is used to describe both concepts. We will have to rely on context to distinguish between the two uses. In this section, I use the word in its general sense, denoting the species, both male and female

[27]I am neglecting the unimportant distinctions between races or families.

Discussion:
The obvious example of relation is fatherhood and sonhood; one cannot have a father without there being a son; there cannot be a son without there being a father. It is impossible to consider the concept of fatherhood without bringing in the son; it is impossible to consider sonship without mentioning the father. Equally employer and employee. Note also, however, that each of these examples involve some sort of hierarchy (subordination is possibly a better word): mother gives birth to the daughter, the father begets the son; the daughter is birthed, the son begotten; but the son does not give birth to the father. One is the the principal or source; the other is the response.

Now, in common English speech, there is another meaning of relation which does not imply hierarchy, for example between brothers. But this in itself does not imply difference. Two brothers are differentiated by their substance, but not by their relationship[28]. Thus in God, there can be no relation of this sort where the two elements are of equal standing in the relationship, as there is no difference of substance within God. If relationship exists within God, it must be in the first sense of a hierarchical relationship.

Now, it is important to stress that this hierarchy does not mean that one hypostasis is worth more than the other. For example, in human relationships an employer often thinks he is worth more than his employee (for example, most will pay themselves a larger salary). Although it is a popular view, I do not accept that it is a proper distinction at the human level: the employer and employee are equal in worth and all things expect their relationship and (perhaps) experience, and all other things (such as their productivity and length of service) being equal ought to be valued the same (for example, an employer and employee of equal experience in their roles, and whose roles require an equal level of technical ability, and who have undertaken an equal amount of personal risk ought to be paid the same, since they have each contributed the same amount of work to the company and both work equally hard for the same goal, namely that the company should succeed in providing whatever services it gives to society as efficiently and with as much quality as possible). Within the Church, the bishop is not superior in worth or value to the priest, nor in importance to the Church; the priest not superior to the layman; the only distinction is how they relate to each other within the temporal ordering of the divine service, and consequently the different roles they are called to play. Within divine relationships, it is certainly not the case that hierarchy implies superiority in value or worth. It merely implies that they relate to each other in different ways and in terms of their origin[29]. □

DEFINITION 13.4.4. **Procession**:
A procession is literally a locomotion from one place to another. In particular, it links the source of the movement (the start of the movement) with the end (or with multiple ends). The source and the end of the procession are distinct (when considered in terms of the procession) in terms of their relationship, in addition to any other differences between them. The word may also be applied analogously to other types of motion or change, and I will chiefly use it in this analogous sense. △

Discussion:
A procession refers to some form of change (though not necessarily change of substance; the essence of a being may incorporate several different states of being, and a change may be a

[28]Again, I neglect for simplicity the concepts of older and younger brothers which is present in this analogy but not in the concept I am using the analogy to illustrate, as this cannot be a distinction within the timeless God.

[29]The word *origin* here needs clarification: it does not imply a commencement in time, but origin of explanation.

movement from one of these states to another. For example, a man may move from standing to sitting down; this involves a change within the man, but whether standing or sitting down the substance of the man is unchanged) or a hierarchical distinction that doesn't involve a change. This may, in some cases, be considered a cause; but not if the change or distinction is within a substance, so the man's standing is not the cause of his sitting down.

For example, we may consider a substance, $|S\rangle$, which is the linear supposition for two states $|X\rangle$ and $|Y\rangle$, so $|S\rangle = a\,|X\rangle + b\,|Y\rangle$ for any allowed values of a and b (constrained only by the condition that the total probability for $|S\rangle$ to be in one state or the other must be one, $|a|^2 + |b|^2 = 1$). Obviously, this picture may be extended to having more than just two states. A procession in the sense of change would be an act where the coefficients a and b are changed. A procession involving distinction may be thought of as a relationship between two elements A and B such that A implies B and B implies A, but which there is still some ordering between them so A is placed as prior to B. For example, to be a musician implies the existence of music, since the definition of musician is 'one who produces music'. One cannot be a musician without music (at some point in one's life); equally one cannot have music without a musician. However, the relationship between these two is hierarchical, since the music originates from the musician. Thus we can say that the music proceeds from the musician. This can be the case even if there is no change in substance involved; for example an eternal musician playing an eternal note. □

CONJECTURE 13.4.5. *There are four types of procession; internal procession, external procession, inward procession and mixed procession.*

Proof:
A particular procession is identified in terms of its source, end, and what causes the relationship between the two. A procession may start from one substance, and have its end within the same substance: an internal procession. It may start from one substance, and have its end wholly outside that substance: an external procession. It may start from outside a substance and have its ends within it: an inward procession (sensation is an example of this). It may start from one substance, and have some ends within that substance and others outside it: a mixed procession. Every procession must clearly be one of these three types, and there are no other alternatives. ■

CONJECTURE 13.4.6. *An act of the intellect involves an internal procession.*

Proof:
An act of understanding requires bringing into knowledge something which was either previously not known, or which was previously known but not considered. (In this context, *previously* does not primarily mean previously in time, but previously in the chain of thought.) It therefore represents a distinction and relation between the knower and the known, and thus a procession. However, the procession is entirely internal to the intellect; the mind may later use the knowledge in its actions, but the 'gaining'[30] of knowledge through reasoning does not directly produce any outside effect. Therefore, an act of the intellect, which is an act of understanding, is an internal procession. The source of the procession is the mind in which the intellect is a component, and its object is the form or other abstract idea being considered; the procession is the link between the two. The mind is distinct from the

[30]Gain is only applicable to a temporal being. A timeless omniscient being does not gain knowledge, but nonetheless the knowledge remains internal to it; in this case the argument remains analogous.

thought (for example, thinking about the form of teapot does not make one a teapot); yet the thought resides within the mind. ■

CONJECTURE 13.4.7. *An act of the will involves a mixed procession.*

Proof:
An act of the will creates a change, and it therefore may rightly be regarded as a procession. It is primarily directed at something outside the mind; it is therefore not an inward procession. However, it also has an effect within the mind. Our actions affect us as well as that which we act upon, not least because before or as we act we will determine to act, which is a process internal to ourselves, perhaps also visualising the outcome. Therefore an act of the will is not wholly an external procession, but a mixed procession. The source of the procession is again the mind that houses the will; the ends are both the external effect and the inward determination to act. ■

CONJECTURE 13.4.8. *The processions of the intellect and will are distinct.*

Proof:
The processions of the intellect and the will are distinct in that they have different objects. For the intellect, the object of the procession is an abstract idea, which is wholly immaterial and must remain immaterial; the will is directed towards an effect in the material world; even the inward change is not concerned about the abstract thought but the material effect. The intellect is directed towards *knowing*, while the will is directed towards *doing*. Equally, in a truly rational agent, the will is inspired by knowledge; one gathers together one's knowledge, and decides how to act on the basis of that. Knowledge is thus, while not the source of the will (which is the mind), an instrument used by the mind in determining the object of the will and thus is related to the direction of the procession and not the object of the procession. Yet knowledge is the object of the intellectual procession. Therefore these two processions are distinct. ■

CONJECTURE 13.4.9. *Procession exists essentially within God.*

Proof:

(1) The procession of the intellect and of the will follow necessarily from the possession of an intellect and will.
(2) God essentially has an intellect and will.
(3) The processions therefore necessarily exist within God.
(4) What is necessary to a substance is part of the essence of that substance.
(5) Therefore procession exists essentially within God.

■

Discussion:
That having an intellect implies the procession of an intellect is clear, since the intellect is that which understands, and the procession marks the process of understanding. One cannot understand without some means by which that understanding is achieved. The only

intellect which lacks a procession is thus that which never attempts to understand anything. However, an intellect is defined by its natural purpose (definition 14.5.2) of understanding. An intellect that never attempts to understand thus violates this natural purpose, and is thus not good (definition 14.4.4). However, God is wholly good (conjecture 14.5.16). Thus, since God possess an intellect, that intellect will involve a procession. A similar argument applies to God's will.

That God has an intellect and will was proved in conjectures 13.3.12 and 13.3.2; these can either exist essentially or accidentally within God, but since God has but the single potentia and an accidental property is one that exists in some potentia and not others, it is clear that these must be essential attributes of God. The rest of the argument is straightforward.

These processions in God indicate distinction and relation, but not change. This is, of course, different from our own intellect and will, because we grasp new knowledge there is a change in our understanding and therefore a change in us. Equally, our intellect is grounded in our brain chemistry and its form, which is in a constant state of flux. It is therefore difficult for us to imagine an unchanging mind. However, the analogy with our own brain fails because God's intellect is not housed in a material brain; and the analogy with our coming to know things fails because God does not learn anything or come to know things through sensation, but instead just has continual perfect knowledge though his act as a sustaining cause. The idea of intellect also does not contradict divine simplicity; we should not confuse the idea that because our mind is housed in a most complex material object that all minds must be housed in such complexity. We should not confuse the simplicity of substance, in this case mind, with the complexity of ideas contained within that mind. Instead, the act of understanding is timeless, but that does not deny the hierarchical relationship between the intellect and the ideas contained within the intellect. That relationship provides the procession. The procession is not temporal or spatial, however, but more akin to the analogy of the musician and the music.

The will is slightly different, since part of the will extends outside God, into the temporal world. Thus there is an element of God's will which is not unchanging, in as much as God's constant purposes adapt to circumstances in time. It is not God that changes, but the events which He observes and directs. There is, however, once again there is the idea of procession between the mind and the object of the will; this procession has to exist in God.

That, of course, leaves open the idea of precisely what God's mind is and how ideas are instituted in it. The problem is that to understand reality we need to map it to some abstract concept, with a one to one mapping between the real object and the abstraction, allowing us to go from the object to the abstraction and back from the abstraction to relate to the real object. This is, of course, how physics works, providing a mapping between physical space-time and the geometrical space-time, and also a mapping between physical potentia and the mathematical notation describing the eigenstates of the Hamiltonian operator, and physical observables with various other operators. We should not assume that this mapping tells us everything about the real physical object and what it is, but it tells us a great deal about it. But we can only understand something in as much as we can create this abstraction. In the case of physics, we can make a start because there are certain aspects of the physical world which share various key properties with the abstraction, such as the topology of physical space and the topology of various geometrical spaces. But with regards to the mind of God, we don't really have a clue where to start. It is not clear that it can be mapped even partially to a mathematical structure in the same way that the physical world can; if it can be it is certainly beyond me. Without this sort of mapping, we have no hope of understanding it. Thus our failure to understand this structure is not surprising: it is an immensely challenging problem, and almost certainly insoluble. Thus at this point I have to say that I don't know because there is little available evidence nor any clear conception of method; the problem may be insoluble, but that should not stop us from trying to search for

a solution, once we have got the easier problems of our own mind and fundamental physics cleared up (there is little point in trying to race against horses when we haven't yet learned how to race against men - or anything beyond the sea-slug). Our difficulty in conceiving such an immaterial mind is related to this lack of available abstraction. That does not mean that the concept of an immaterial mind is incoherent. It merely means that as yet our knowledge is very limited. Indeed, the concept must be coherent; because while we cannot say what an immaterial mind is, we can say that it is, using the arguments above, or similar arguments, or by appealing to observation (i.e. revelation).

Logically one cannot have a speaker without having a word (whether the speech is internal or external), as the definition of a speaker is that which speaks, which speech necessarily producing a word; if one does not produce speech then one is not a speaker. Equally one cannot have a word without a speaker of that word; for a word is that which is spoken. Thus the speaker and the word are not separable if there are no other considerations, and especially so for the internal voice (for the external voice, the word proceeds through another medium such as air, and thus no longer remains directly connected to the speaker except as an efficient cause). Thus the speaker and the word of the internal voice (which is the object of the intellectual procession) are of one substance. The distinction between internal word and the origin of that word is not one of different substance, but a distinction within a substance, and this does not contradict the proof of divine simplicity. Equally, the thought in the divine substance does not entail change, and thus does not contradict the proof of divine immutability. □

CONJECTURE 13.4.10. *There are two distinct processions within God.*

Proof:

(1) Of the mental attributes, the mind of God only contains intellect, will and freedom essentially.
(2) There is a procession of the intellect.
(3) There is a procession of the will.
(4) There is no procession related to freedom.
(5) The procession of the intellect and the will are distinct.
(6) Therefore there are two distinct processions within God.

■

CONJECTURE 13.4.11. *The end of these processions are distinct from the source.*

Proof:
There are logically three possibilities: the end of a procession may (in principle) be distinct from the source of the procession; or it may be identical to the source of the procession, but at a different time or place (or any similar mark of difference; so, for example, a thought may be emitted from the mind and the resultant change be registered in the same mind but at a slightly later time); or it may be identical to the source of the procession at the same time and place (and everything else) as the source. The third of these is impossible, since a procession implies movement from somewhere to somewhere else. The second is impossible within God, since, as shown above, there are no distinctions in terms of time, space or other marks within God. Thus we are left to conclude that there is a distinction between the end

of these processions within God and their source. ■

CONJECTURE 13.4.12. *There exist three hypostasis within God.*

Proof:
The only point left to prove is whether there are three or four divine hypostasis; that is whether the sources of the processions should be identified (three hypostasis) or kept distinct (four hypostasis). But if they were distinct, there would be no connection between the procession of the intellect and of the will; but this is not so, since the will is guided by the intellect. There are thus three hypostasis within God, and not four. ■

CONJECTURE 13.4.13. *These hypostasis may be referred to as persons.*

Proof:
A person is defined as an individual element (which includes the concept of hypostasis) of a rational nature (definition 13.4.2); with the rational nature referring to the highest genus where rationality was an essential element of that genus. Now, God is unique; so the primary distinction within God is that of the source and ends of the processions of the intellect and the will. Also God is essentially a rational agent. Thus these hypostasis are rightly referred to as persons. The word *person* here is primarily used as distinguishing between the different elements of God's nature. ■

Discussion:
Now it may be objected that we ourselves have an intellect and will, and therefore processions of the intellect and the will? Does that mean that we ourselves are examples of three persons within one substance? However, several elements of the above arguments do not apply to us. Firstly, we are members of the human species, and it is the essence of our species to be of a rational nature, while we inherit that essential quantity from our species. The word person rightly distinguishes between different members of the species, rather than distinctions or differences within an individual. Secondly, in the above argument, we have made use of God's timelessness, which does not apply to us. It may be that the object of the procession is the source, but at a different time. Equally, we have implicitly made use of God's simplicity, which again does not apply to us (at least, we have not shown that it does). Thus a procession within God's mind must be within a single substance; however if the human mind is complex (it has both an immaterial aspect, but also originates within the material brain, itself a complex structure) then the discussion of internal and external processions is complicated.

Nonetheless, we do possess processions of the intellect and the will. They are not exhibited within us in the same way that they are within God. The processions within us resemble the divine processions analogously rather than identically. It can be used as a picture to help us to understand God, but it remains only a picture and therefore will be different from the processions within God in important aspects. □

CONJECTURE 13.4.14. *The persons are distinguished by relationship rather than substance.*

Proof:
Different objects may be distinguished either by substance, relationship, or both. This is clear; if two items are of the same substance (which implies the same matter and the same form for material objects) and are yet distinct, then they must differ by relationship; that is there is some non-symmetric connection linking them (and two objects with the same matter must be linked by something) which distinguishes them (for example, one came first in time and the other later; *earlier* and *later* are statements of relationship as in this case the connection between the objects is one of time[31].) Now the divine persons are not distinguished by substance. Therefore they are distinguished by relationship. ■

CONJECTURE 13.4.15. *The divine procession of the intellect may be referred analogously as begetting.*

Proof:
Begetting (for example, the concept of fathering a child) can be understood in two senses. Firstly, it implies the origin of the child. But not everything that comes into existence does so by being begotten. Therefore, the primary sense of begetting refers to similitude, in particular a similitude of form. A man makes a table, which has a different form to him, but begets a son, who has a similar form, in that both father and child are of the same species. Now the intellectual procession is also marked by similitude, as both the origin and source are within the same substance. Therefore, we may refer to this relationship as begetting, and the origin of the procession analogously as 'Father' and the end as 'Son'. ■

Discussion:
The names 'Father' and 'Son' are given analogously rather than identically; a human father and son are of similar substance (they are not identical in form, although they share a great deal in common), while the divine Father and Son are of the same substance. However, they are appropriate in that they stress that the primary difference between the persons is that of relationship rather than a difference in form. It is also appropriate to refer to the Son as the Word of God or the Wisdom of God, signifying that the relationship is the procession of the intellect.

The name for the third person of the Trinity is less clear, since there is no analogous concept within human relationships. I shall therefore take the name 'Holy Spirit' from Christian revelation. The name is appropriate, since *holy* means set apart for God, and *spirit* refers to the manifestation of an immaterial aspect of a mind. The Holy Spirit is, in Christian theology, that element of God through which the Christian is united to Him. It is, therefore, an aspect of God which is directed in part towards that which is external to him. While we have so far considered the inward aspect of the procession of the will, it also contains an outward aspect, which confirms that the idea of the Holy Spirit not only being related internally to the Father and the Son, but also externally towards the material world, including ourselves. Equally, there is no easy analogy we can use to describe the procession of the third person (as we use begetting or generation to describe the relationship between the Father and the Son), so we simply use the word proceeds (for example, in the Nicaea/Constantinople creed, in its original form[32], we recite that the Holy Spirit is 'proceeding from the Father,' while the Son is 'eternally begotten of the Father.'

[31] Two different objects of the same form cannot occupy the same matter at the same time; if they did they would be the same object.

[32] I do not wish to comment either in favour of or against the *filoque* here; my own position should be clear since I accept the classical Anglican formularies.

It may also be questioned why we use 'Father and Son' and not (say) 'Mother and Daughter.' The main reason for this is that the revealed character of God is more masculine than feminine: God relates to mankind more as a husband relates to a wife than as a wife relates to her husband. The relationship between husband and wife is not, of course, symmetrical, as evidenced by their different roles in the family and in particular with respect to the children. The different physical and mental characteristics of men and women are a manifestation of this difference in relationship. This relationship between men and women is not one of superiority and inferiority, but of relationship and role particularly within the family (but not only within the family). Thus using the word Father to describe God is better because it brings to mind the dominant aspects of God's character and the nature of His relationship with us. However, this truth concerning God is taken from revelation rather than from reason, so I do not wish to dwell upon it in this work. □

CONJECTURE 13.4.16. *Each of the divine persons shares in the divine attributes, such as timelessness, omnipresence, majesty, omnipotence, and so on.*

Proof:
The divine attributes, as shown above for those attributes considered here, are essential features of the substance of God. One cannot be of the divine substance without possessing the divine attributes. Yet the three persons of the trinity are all of the divine substance. Therefore, they must all possess these attributes. ■

Discussion:
Note that we regard the persons of the Trinity as being equal in the divine nature but not identical[33]. They equally share the divine attributes, but are distinct in both their relationship to each other, and their relationship to us. In particular, the Father is the origin of the Son, and this is the primary distinction between the Father and the Son; the Father is the Subject of the verb, and the Son the object. So the Father sends[34], the Son is sent; the Father gives[35], and the Son is Given; the Father directs and the Son obeys[36]. Thus, within the Trinity there is both equality of substance and attributes but hierarchy in relationship. That the Son does as the Father commands does not imply that the Father is superior to the Son, but is merely a description of how they rightfully are meant to relate to each other. For example, suppose that there are two identical twins travelling by car to a particular destination. One must drive, and the other navigates. Which of them takes which role is up to them, but once the role is taken they must stick to it; if they each try to go by their own preferred route there will only be confusion. That the navigator gives directions and the driver follows them does not imply that the navigator is superior to the driver; it merely describes their relationship. The navigator has the skill of reading maps; the driver the skills related to driving; and both are essential if they are to reach their destination. If there were two drivers or two navigators, they would not get there. Now this analogy is flawed in numerous ways, most importantly in that the navigator and driver voluntarily choose their roles, so they take them on accidentally, while the relationship between the Father and the Son is essential, and that the navigator and driver are differentiated by function while the Father and Son are distinguished by relationship; nonetheless it does illustrate how subordination does not imply inferiority.

---

[33]The difference between equality and identity is that equality implies being like in certain respects, while identity means being indistinguishable; like in every respect.

[34]John 20:21.

[35]John 3:16.

[36]John 14:36.

There are two crucially important but difficult aspects of Christian theology, the doctrine of the Trinity (God is three persons of one substance) and the doctrines regarding Christology, the person of Jesus Christ (the Word of God taking on human flesh; both fully divine and fully human with the natures inseparable). Since this work is intended to concentrate on classical theism rather than Christianity, I will not discuss Christology here; but explanations of the doctrine can be found in the various texts of philosophical theology, not least Aquinas' *Summa Theologica.* I will not claim that anyone, least of all myself, fully understands either of these two doctrines. That does not mean, of course, that they are not understood partially or incompletely. Numerous models of the Trinity have been proposed throughout the history of the Church, including the model presented above (which is not original to myself, but adapted from the medieval theologians, who in turn adapted it from the classical theologians). The question is, I think, more a matter of a lack of evidence to guide us than a lack of our ability to understand.

The argument above seems to suggest that in order to think, there must be some structure within God. This need not be as complex as the human brain; in fact it cannot be given God's immaterial nature and simplicity. Simplicity of substance, of course, does not imply that everything related to that substance is simple; there is a difference between the substance of a mind being simple and the knowledge and ideas within that mind being simple. Consciousness arises in us in one way, but it need not arise in everything that is conscious in the same way; and it need not be experienced by everything conscious in the same way that we experience it.

Complexity often arises from simple forms in physics. An example I am most familiar myself with is quantum electrodynamics, where the Lagrangian is very simple and straightforward (at least, once you understand the mathematics) – in fact, it is just about the simplest non-trivial Lagrangian you can have in four dimensions; but it does not mean that the physics arising from that Lagrangian is simple. Indeed, all of chemistry is contained within that one short line of algebra; the forms of the molecules and their possible interactions (their efficient and final causes) hidden but implicitly described within the Lagrangian. There are various models studied by physicists – the Ising model, the Schwinger model (quantum electrodynamics in one space and one time dimension), the Ginsparg-Landau model; where a great deal of structure necessarily emerges from little input: the number of dimensions, the symmetries, and the types of particles. From this, we find different phases of matter with very different properties (think ice, water and steam) emerge by just altering the parameters a little. Simplicity of substance (described by the eigenstates of the Hamiltonian, or equivalently the Lagrangian, for material objects) does not mean simplicity of effect.

For example, a Hamiltonian may have several distinct ground (lowest energy) states. We can consider the ground state of a quantum field theory with a Z_N symmetry, where there are N ground states evenly spaced around a circle. The Hamiltonian possess a symmetry, in this case a rotational symmetry, which maps these ground states to each other. But nature has to choose one of these ground states; the symmetry in the Hamiltonian is broken in nature; the simplicity of the Hamiltonian is broken necessarily (this is known as spontaneous symmetry breaking). Now let us suppose that each of these ground states represents a different bit of information; a letter. Then by building up a sequence of these systems, we can convey any information we choose just as by building up a sequence of letters, we can write a book. The Hamiltonian (and thus substance) is relatively simple; yet it can describe information of infinite complexity. But storing information requires some multiplicity of ground states; the repetition of a single letter cannot convey any meaning – we need at least two letters. In the same way, I would suggest (although, of course, I cannot prove) that if knowledge is contained within God there must be some structure to differentiate different letters (or whatever passes for letters in the divine thought), and if there is some analogue of spontaneous symmetry breaking this need not conflict with the simplicity of God's substance. Granted this picture is incomplete; one must have a way of

interpreting and making use of the information; just as the information contained within DNA molecules is 'read' by the various chemical reactions which take place within a cell. Even more importantly, any picture based on a material system must fall well short of describing the nature of an immaterial God; and the Z_N group is not perfectly simple (which would be represented in physics by having the maximum possible symmetry). Nonetheless, this picture does illustrate how complex information can emerge within a simple substance. □

CHAPTER 14

Ethics

14.1. Introduction

One of the most common objections to theism is the problem of evil: how could a single, benevolent, omnipotent, God create and sustain a world that contains evil? I would be remiss if I did not address this issue. However, even to pose the question requires more work than is often supposed and the question contains numerous presuppositions and assumptions that need to be clarified before it can be asked, let alone answered. First of all, we need to ask ourselves what does 'benevolent' mean; what does 'evil' mean; and most of all how are we to define moral goodness objectively, a task which many modern philosophers claim is impossible to do? This leads us into the murky ground of ethics. The issue here is that there is not one universally agreed upon ethical paradigm. The answer (if there is an answer) to the question depends on the underlying ethical assumptions; what the question is asking similarly depends on these assumptions. Thus before we attempt to answer the questions of whether or not God is good, we need to study the question of goodness, and moral goodness in particular.

But doesn't everyone intuitively know what is good and what is evil? Isn't this enough to get at least some common ground? No. For example, almost everyone would agree that it is wrong to commit robbery and murder (though even here there is confusion: what someone describes as murder another might call a just act of retribution to preserve the family's honour, as one of many examples). But the problem of evil demands that we go beyond these intuitive understandings, and to a truly objective and rational treatment of the subject. Why? Because it is at the most fundamental level we find the link between ethics and God; ethics is derived from metaphysics, as is our (rational) understanding of God. This link through metaphysics between ethics and theism is crucial; for example to maintain self-consistency we need to use the same philosophical basis to define both 'evil' and 'God' in the question.

The problem of evil is essentially an argument from contradiction; it depends on various premises one of which is the definition of God and another the definition of evil. From these definitions we derive an understanding of how we expect a good God to act in the world; from the definition of evil and observation we obtain a picture of how the world really is. We then note that the theoretical understanding contradicts the observation; and conclude that some element of the argument is thus in error. There is a contradiction somewhere, and those atheists who wield this argument would claim that it is in the notion of 'God.' Which is all very well, but it goes beyond what has been proved: there are more places where the chain of reasoning could break down than just that. For example, if our understanding of what is evil arises from different assumptions from our understanding of God, then we immediately have the problem -- there is no need to question the notion of divinity, because the argument is built on contradictory premises, which has nothing to do with the existence of God. Equally, if we derive a notion of 'God' from the premises used to build up our preferred ethical system, then the argument might be sound, then all we have shown is that that notion of God – which might well be one that nobody in practice worships – is incoherent; and we have done nothing towards disproving the God that many people believe in. Rather, we have to construct our ethical theory from the same premises that give rise

to an idea of God which aligns with some major religion; for example the premises behind classical theism. If one could then show a contradiction between the expectations of that ethical/theological theory and observation, then the theist would have a problem. But, of course, to do that means working through ethics in detail from the fundamentals upwards.

The first and most important thing to note is that an ethical theory depends on the underlying metaphysics: it has to, because there is nothing else it can depend on, at least if we are to come up with an objective answer. This is, of course, well known: both Hume has argued that it is impossible to get from an 'is' to an 'ought', and Moore has argued along similar lines for the undefinability of the concept of goodness (meaning that we can define some rule which states that 'X' is good, and we can say that goodness is what we ought to be, but these are, in fact, two different definitions of the word 'good' and, according to Moore, there is no way to connect one to the other). But these arguments are only valid if their premises are valid. Hume, of course, threw metaphysics into the flames, so one of his premises is that we cannot use metaphysics to provide the link. Hume's conclusion, then, should be stated as we cannot go from 'is' to 'ought' while adopting Hume's own (lack of) metaphysical principles. Moore was better, in that he mentioned metaphysics in his argument for his 'naturalistic fallacy', but also failed because he did not consider metaphysics in general, but only certain approaches to it. If we are to believe these two scholars, then, the conclusion should be that an approach to ethics that is not founded on metaphysical principle cannot be objective. And since ethics is meant to give us universal rules of behaviour, how do we say that our belief that 'X is good' is good, without some objective principle to guide us? Clearly we will be left with either an insoluble dilemma or an infinite regress.

There are two sides to ethics: 1) Constructing a general ethical framework; 2) applying it in practical situations. Here I solely consider the first of these: I am interested in general principles rather than specific applications, adjusted according to each circumstance. This does not mean that I consider the applications unimportant, nor that the system that I advocate is great in theory but fails when it comes to the practical details. My consideration is solely that a) In the order of understanding, a general framework comes first, with the applications derived from it plus a knowledge of the local circumstance, so even if I were to discuss practical applications I would still have to start by establishing the wider principles; b) This is not a full treatise on ethics, and I only intend to take what I need for my own purposes and then move on without being bogged down by unnecessary digressions; c) given the practically infinite number of particular circumstances most of which I am unfamiliar with, even if I were to discuss the more practical side of ethics, I could hardly hope to do it justice; and d) better people that I have already done that work (for example (Oderberg, 2000; MacIntyre, 1999; Oderberg and Chappell, 2008; MacIntyre, 2016)). More practical applications can be derived from the framework I describe here; I just leave it to the reader to do so for their own circumstances.

Before we begin, I ought to define what I mean by ethics; and here we encounter our first subtlety.

Consider the following definitions,

DEFINITION 14.1.1. **Deed ethics:**
Ethics is that branch of philosophy which studies how rational agents ought to behave. Δ

DEFINITION 14.1.2. **Being ethics:**
Ethics is that branch of philosophy which studies what rational agents ought to be. Δ

That it is a branch of philosophy means, of course, that it is to be studied using the tools of philosophy; namely logic built on premises which are preferably either tautological, empirical or the axioms of logic.

The two definitions are, of course, related. If we take the second definition, for example, then we immediately get a description of actions; for example, a compassionate action may be defined as that which compassionate people tend to do. Whereas, if we take the first definition, we may think that we get the second definition; someone who routinely acts in a compassionate manner has the appearance of being compassionate. Most of the time, appearance will correspond to reality; but this need not be the case always. For example, one can act to help people take out their garbage because one is paid for it (a selfish and non-compassionate motive), or solely because of a desire to help them (even if one is later given a small amount of money). The action in both cases is indistinguishable; the motive is not. Of course, in real life people may act from a mixture of these two motives; but let us consider two people, Albert, who acts for money and if he were not being paid would scatter the garbage across the living room floor, and Bill, who acts solely out of a desire to assist and would return the money if he did not know that the person he is working for would insist on his taking it. On the first definition of ethics, both are acting morally (if taking out the garbage rather than throwing it on the floor is judged as being good), since both act in precisely the same way; in terms of their actions, there is no way to distinguish between Albert and Bill in this instance. Using the second definition, only Bill is ethical, since only his actions emerge from a good character. Thus these definitions, while similar, are not equivalent. They are related to two of the schools of ethical thought; definition 14.1.2 leads naturally to virtue ethics, while 14.1.1 to duty-based ethics and, to a lesser extent, consequentialism.

Note that definition 14.1.2 most properly describes the formal cause of the rational agent (which relates to what the agent is), while 14.1.1 focuses on actions (what the agent does) which come under the purview of efficient and final causality. Now enlightenment philosophy denies the existence of formal and final causes; it is therefore, to be consistent, limited to definition 14.1.1, and I imagine that this is what most people today think of when they think of ethics. To the Aristotelian, however, definition 14.1.2 is more fundamental and of the two ought to be preferred.

There is, however, one further problem with both of these definitions, and that is the word *ought*. What do we mean by this word? The answer, I think, depends on one's school of thought, which ultimately means that the definition itself brings in various metaphysical assumptions. People with different world-views can't even agree on what *ethics* means.

I personally prefer a third definition, which avoids the issue of what we mean by *ought*,

DEFINITION 14.1.3. **Ethics**:
Ethics is that branch of philosophy which studies goodness, and its applications to the lives of rational agents. Δ

Now this obviously presents another issue, namely what we mean by *goodness*, and I shall define this later. This definition is more fundamental than the previous two, since our conception of *ought* should be constructed from our concept of *the good*. Since the concept of goodness is more fundamental, we should keep it foremost in our mind and therefore I believe that this definition should be preferred.

If the subject of ethics is the status of men as moral agents, then an important question is what is the moral status of man. We know that all people are a mixture of good and bad. How does this come about? Consider the following two views:

(1) Man is born free, but everywhere he is in chains. The natural end of man is liberty. Although intrinsically free, man is held back by those around him who pressure him to conform to ill-conceived biases: these bonds on him may come from the

family, the global patriarchal conspiracy, the local community, pressures to accept other people's sense of your own identity, the church, historically dominant views or the bourgeois classes: the instruments of inequality and prejudice, where men who are by natural right equal are treated as subordinates and therefore as inferior. Through these instruments, he loses his birthright and becomes trapped in a life of misery, later propagating others to the same position because he knows no better. Value is given only by the labour put into product, but this labour is stolen from the workers by the hierarchy, to the benefit of the few private individuals who gain from the oppression of the masses. The consequence of this are the power struggles and evil we see in society. To become pure, we must rise above these bonds; and the instrument for doing so must be impersonal and above those structures which hold man back: an egalitarian (political) state, which emerges when formally oppressed groups and societies rise up to take back the liberty that was stolen from them. Because all evil comes from the struggle for power over other social classes and victim groups, the possibility of class, hierarchy and national identities must be removed from society and absolute equality imposed. The governing classes, being less restricted, more educated and enlightened (once the unenlightened usurpers are cast out by the natural processes of progress) and therefore more able to guide society, are least affected by the negative effects of the bonds of local prejudices, and therefore can aid the 'little man' most effectively. The role of government should therefore be expanded to promote the common and greater good, and fight against and outlaw the prejudices and inequalities inherent in the smaller structures of family, community, business, schools and the church, eventually destroying them or incorporating them into itself, so that, through careful central planning, regulation and control to eliminate the oppressive and corrupting class structures, Man, once free of his bonds, can all rise as equals to fulfil their naturally good state. Once this is achieved, and a society of true equals is born, then government itself will fade away as its function is fulfilled in this time of universal prosperity and peace.

(2) Man is born a slave to wickedness, though some have gained a small semblance of freedom. The natural state of man is to act selfishly; and his life is a constant battle against this tendency towards selfishness in favour of what his intellect, if working properly, tells him is a better way: righteousness and virtue. However, even his intellect often does not work properly, and if left to itself will be blinded by his corrupt nature to arrogantly try to justify and support his evil. His first tutors in virtue are those around him: his family, community, church and the historical wisdom of his ancestors. By forming relationships with these people he learns empathy and compassion. These can train him to recognize, and at least partially overcome, his corrupt nature, and to nurture the virtues which allow him to live in harmony with the community. The local community will come together to aid the poor and marginalised among them – only this personal touch can respond to diverse individual needs. As people mature, they learn to trade their labour to supply what someone else needs, with every trade between individual private citizens an exchange of mutual benefit. Men are of equal worth, but are not equal in every respect but instead have different gifts and talents, and correspondingly should have different roles in society; but the local community needs everyone. Those less able should be aided by the more able around them, and each learn to use the skills they have to build up the community, always being aware of the risk of the selfish nature re-emerging in each individual. Government is a necessary evil, needed solely to provide those services which cannot be efficiently organised or financed at a local level, and to support and enhance the local institutions. Decisions are made at as local a level as possible, preferably by those who personally desire a good and service and those who produce it. Since things

> are only worth what someone is willing to pay, the value of a good or service can only be decided by the free market negotiations between the individual buyer and individual seller. Only these two, relating and emphasising personally, can choose a fair price that benefits both parties. Virtue is only encouraged by accountability and individual relationship. Those who lack proper accountability are particularly prone to reverting, at least in part, to their natural state of self-justifying wickedness. These are those who have the least connection to the people whose lives are affected by their decisions, or who can impose laws, regulations or constraints on others without any personal cost. That corruption will make them tend to enrich themselves and their friends at the expense of the people they don't encounter, often without them realising what they are doing. Such people are to be found most of all in government or large publicly owned corporations, or effective monopolies. Their ignorance of the people their laws and decisions affect will lead to the decisions of the central planners being ineffective or counter-productive. Their bad choices will increase the poverty and inequality of society, even if (and perhaps especially if) they rose to power on the promise to protect the marginalised. The corruption that comes from centralisation will lead to an increasingly unequal and immoral society. Thus the powers of government, trade unions, mega-corporations and other national organisations, which attempt to take authority away from the many and into the hands of the few who manage to gain control of such societies, should be limited (and, as much as it is possible, eliminated), and business (of which small and medium-sized local businesses should be preferred where possible - so every decision can be made by the individuals directly affected), academia and the media be regulated by free and fair competition, avoiding cartels. This competition controls the worst of the vices of man. Everything should be done to nourish families, churches and local communities, and encourage them to be as self-sufficient as possible, since only small interdependent societies such as this where everyone knows, cares for, and is looking out for everyone else, can correctly train people in the virtues.

Both of these positions are (or would be, once fully worked out) rational, in the sense the premises lead to the conclusions; yet both give an opposite idea about how society should be ordered. Starting with different ethical perspectives, we wind up in very different 'ideal' societies; the first the more left-wing view of a strong government guided by an expert elite to enact a centrally planned and organised regulation of the groups and communities in society; the second the right-wing position of self-regulating local communities consisting of individuals with complementary skills bound by close personal relationships and guided by a strong virtue-based moral code with a focus on evidence rather than 'expert' opinion, limiting the power and need for central government to the minimum possible.

To discover which perspective is more accurate, we need to compare each conclusion and premise with observation. Whether those at the top of a centralised society tend to spread wealth to all equally, or hoard it for themselves and their accomplices. Whether those nations which favour centralisation, prioritise groups over individuals, and adopt a labour theory of value have been more prosperous than those which promote localisation, individual freedom, and adopt a subjective theory of value through the free market.

But, of course what I have set up here are just two extremes, and our choice is not just between these two positions, but must consider the various positions between them and in different directions.

My own perspective of man is slightly different from either of these views, as I will outline later. The two world-views I outlined above are, of course, as I have presented them over-simplified straw-men; I am not suggesting that anybody holds them as they stand, and those few who stand in broad agreement with either one of these positions would no doubt express them differently and far more clearly than I have done here, without the

obvious flaws and limitations, and adding such elements required to make them complete and consistent. I also do not claim that these two systems are comprehensive; they are both built on the Western post-Christian way of thinking; and would both appear quite alien to the communal honour and religion centred societies of the Islamic world, or the hierarchy-centred world-view of the Confucians and related Chinese philosophies.

Please note that I use the word 'evil' in a sense which is different from the modern world. When the modern world speaks of evil, it thinks of an extreme, such as the deeds of Hitler, Stalin or Pol Pot. However, like many who admire classical philosophy, I use the word *evil* to describe any deviation from goodness, no matter how slight; the smallest lie, lustful glance, or deliberate inhaling of cannabis is an act of evil. The reason for this is to emphasise the severity of such actions; lesser words such as *bad* or *mistake* encourage people to think that they can be excused, or do not matter much. To my mind, these excuses are badly mistaken. Of course there are degrees of evil; I am not claiming that letting your thoughts wander when a pretty girl passes you in the street is as bad as genocide (or could be used as a valid excuse for genocide). It is, however, nonetheless a deviation from goodness; and thus an evil.

It is clear fact that men sometimes act for good and sometimes for evil. Some explanation for this curiosity is required. The two perspectives outlined above take the two over-generalised extremes of the range of possible explanations; one is that man is born pure good and trained by community-wide oppression of other communities to be evil (though some spark of the original good always remains); the other that man is naturally pure evil and is trained by individual relationships to partially overcome this (though some spark of the original evil always remains). Most people would, I think, rightly say that reality lies somewhere between the two views. But before we can answer this, we need to know what evil is, and that means knowing what goodness is.

CONJECTURE 14.1.4. ***An ethical theory, to be useful, must provide a universal, objective and rational standard.***

Discussion:

By rational, I mean that it is self-consistent, accessible to reason, and does not change. 'Accessible to reason' does not necessarily mean that it is wholly derivable from rational arguments, but that it is based on a small tangible number of premises which, ideally, follow from definition, tautology or observation, and given those premises everything else follows and makes sense. In particular, the standard for ethics cannot be based on arbitrary whim: we cannot (in a duty based ethics) demand 'do this' one day, and then the next, in exactly the same circumstances, demand 'do that'. Such an ethics is not accessible to reason. The reason we desire that ethics should be rational is firstly that it is a branch of philosophy, and as such reason must be our chief tool when pursuing it, and secondly from our assumption that the universe as a whole should be rational.

By objective, I mean that it is independent of the human will; no one person can choose for themselves their own standard; no small group of people can choose the standard for others. It should be deduced (as best we can) from true observations about human nature (whether made directly, or revealed by something outside humanity, such as God), not on any ideology about what human nature is wished to be. The reason is simple: an ethical theory is designed to be binding on either the will or character. But this is not possible if the will or character decide on which ethical theory to follow for themselves. We can, of course, set our own targets, and there is much merit in doing that, but such a process does not account for moral truth. If we can set ourselves targets, then we can also change them. But based on what? Our own prejudices or desires? But ethics is there to regulate our prejudices and desires, and tell us what is good and what is ill. Something external to us? Then that external thing is either an objective premise, or the authority of someone

else. But, if it is the authority of someone else, then the difficulty is not resolved, but only repeated at a higher level. And what guarantee do we have that any standard we adopt for ourselves will be good? If ethics is to teach us anything, we cannot be the ones to determine what it is[1]. It is the examiner who sets the pass mark, not the student.

By universal, I mean that it applies to everyone. Again, this can hardly be otherwise: if we say that one set of standards applies to one group of people and another to another group in a different circumstance, then those two statements together, and the principle which inspires both of them, give us a rule which applies to everyone.

Finally, by useful I mean that the ethical theory is of some benefit to us. That is, it can provide us with accurate guidance either about what we should do (in a duty based system) or what we should be (in a virtue based system), and that such guidance should be possible to understand and to follow.

Now, we have not established whether such a system of morality exists: that will be one of our tasks. The conjecture was not that such a system exists, but that only that only such a system can be useful. It may be (at this stage of our inquiry) that ultimately there is no ethical rule, that all talk of 'good' and 'evil' is meaningless; we ought to do just what we want to do. Indeed, some philosophers, most particularly Nietzsche, have, based on a particular set of premises, argued that this is precisely the case: that 'good' and 'evil' are only what the rulers of society claim that they are. Of course, this conclusion is only as good as its premises; and if we cannot accept the conclusion then it is perfectly logical to do so by dismissing the premises. For some, Nietzsche is convincing; for others, his conclusion provides a *reducto ad absurdum* to show the follies of his whole line of thought.

Equally, some may object that 'universal' is too strong; that it implies the same thing to all people regardless of circumstance or character. This, however, is not so. A universal rule allows for differences in circumstance. For example, the virtue of courage is traditionally seen as lying somewhere between cowardice and recklessness. So sometimes to rush into battle is the fruit of courage, and other times (when the odds against success in battle are overwhelming, while there is a much better chance if we fight the battle tomorrow when our reinforcements arrive) to fight the battle is against courage, but is reckless and foolhardy. Similarly, there are times when to withdraw from battle to negotiate a surrender is from cowardice; and other situations when it is the courageous thing to do. Courage is universal; what it leads to in practice depends on the situation. A rule is universal when it covers every situation, not because it demands the same thing in different situations.

Equally, does this proposition mean that every man has precisely the same ultimate standard against which they are judged? No. The standard is based on human nature. There are aspects of human nature which every human being shares; and therefore there may well be, on account of these similarities, parts of ethics which apply to everybody. But there are also differences between people; perhaps most importantly the differences between men and women, but also differences between individual men and individual women. So there is no reason to say that what is good for a man would also be good for a woman. Most of what is good for a man is likely to also be good for women, for we are far more alike than we are different; but there is no reason to suppose that those aspects of the ethical standard which are derived from those aspects of human nature where men and women differ will be

[1]Some may object that no standard is objective; after all everything has to be evaluated, judged and decided on by people, who each bring their own biases and presumptions. But there is a difference between saying that there is an objective standard which we try to emulate (even if we fall short), and something which is purely subjective. There will, as in any scientific endeavour, be imprecision in places in our ethical theory; where we can not say 'do this', but rather 'it is either this, or this, but it certainly not that'. We should acknowledge where those uncertainties lie, and neither under- nor over-estimate the errors. The task of philosophers is to reduce those errors as much as possible.

To deny the possibility of an objective truth is foolish. Physics is not invented by the arbitrary whim of the physicist, but the physicist is guided by experiment and constrained by reality, and ultimately has no choice about his destination. It ought to be the same for ethics; otherwise ethics is useless.

identical for men and women. Equally, it is accidental to human nature whether one should teach history at a university, or be a professional sportsman. It is good for the university professor to spend most of his time reading and writing books relating to his speciality; but bad for the professional athlete, who ought to concentrate instead on his physical training. A sportsman may become an world-leading expert on medieval history, but it will hardly help him win an Olympic medal. However, for both of them, idleness is a vice; the lazy man does not win a chair at a university; the athlete who skips his training does not win the prize.

In studying ethics, we are first interested in those virtues which are universal to all men. It is necessary to master the general principles before applying them to specifics. This is not to say that there are not other additional principles which are particular to specific people: but the general principles come first, and the specific principles must be made to conform to them. There must, of course, be no contradiction. □

Objection:
Why bother with all this philosophising? Is not our moral intuition enough? Surely, as long as we don't harm others, we should be free to do what we like? •

Response:
There are numerous difficulties with this proposition. Firstly, the concept of 'harm' is undefined. We might think that we intuitively understand what harm is, but do we truly do so? Different people have different opinions. Which of us is right? Are any of us right? What if to avoid doing what the person we affect considers harmful, we have to do what we ourselves consider harmful? Should we be bound by our conscience, or his? Everybody makes mistakes. What we might think of as harmful may in fact be beneficial; what we think of as beneficial may in reality be harmful. Can we use what each person believes to be harmful to themselves as our definition of what it is to harm them? But what that person believes, even about himself, might be mistaken. A student might resent extra mathematics lessons at the time, but ultimately he will gain from them. An athlete might reject the extra training that his coach gives him, but avoiding it would in reality be more harmful to him than doing it. Can we define harm as what the person thinks when he is sober and in his right mind, and truly knows the effects of the action? But are any of us ever in a right enough state of mind to make such judgements, and how can we be sure what somebody would think at a later time? Maybe the student would regret the extra lessons even after he has enjoyed the benefit of them, still believing that a night of wild partying would have served him better.

Secondly (and related), why restrict ourselves to avoiding harm? There is the obvious issue that we should also seek to benefit others, but, more importantly, why suppose that causing harm is always morally wrong? Of course, we might define harm as *that which is morally bad*, but then we have the opposite problem that we have redefined the term *harm*. It no longer means what our intuitions, say, but something which is morally bad. Which means that we require a definition of moral badness that doesn't rely on the concept of harm. Assuming we don't do this, we are left with a deeper issue: why should we assume, without proof, that to harm someone is *always* a moral evil? I am not saying that it isn't; only that unless it is proven, the statement could be wrong.

Thirdly, why restrict ourselves to harm (and benefit) to others? Should we not also think about the harm done to ourselves by our rash actions?

But fourthly, and most importantly, this view begs the question. Why should we be free to do anything? Surely, a purpose of ethics is to determine what we should be free to do, and in what aspects we ought not to be free. That we are free to do something does not

mean that it is good for us to do it. We might, after going through all the work, determine that we are free to do anything, but that should be the conclusion, not the starting point.

Our intuitive moral codes might be a start, but like all intuition, they are no substitute for rigorous thought. Our intuitions are shaped from our culture and environment, and what is our culture and environment but simply the prejudices of our own age? There is no guarantee before we start thinking about it that our culture has things right, and the culture of our ancestors or another nation were wrong, any more than there is that our ancestors were wrong and we right. This humility about our own age is particularly important because our ancestors in general tended to think about moral questions far more rigorously than most people do today (except, perhaps, those trained in philosophy – though even there I am uncertain that today's philosophers are a match for their predecessors). Maybe we know more today than was known in the past, but have we taken advantage of that knowledge? And what have we forgotten? And which errors have we picked up that we ought to forget?

That is not to say that our culture's moral intuitions are wrong. They might be, they might not be. But it is to say that if we are to find our moral bearings, we need to tackle the problem rationally and rigorously. And that means using logic based on sound metaphysical and observational principles. We cannot avoid the hard work with an appeal to intuition. ○

14.2. Types of Ethical Theories

There are three principal types of ethical theory, just as there are three parts to a sentence: subject, verb and object. A virtue based ethics looks at the subject, and asks as his starting point what sort of person someone ought to be. He will then later move on to actions and consequences; but by asking (if courage is a virtue) what a brave man would do in that circumstance. It is not the action in itself which is good or bad, but the motivation and character of the one who performs the action. The duty based ethicist, on the contrary, sees actions in themselves as being right or wrong. To lie is wrong, regardless of the motivation or consequences. The consequentialist (of which the utilitarians are the most prominent) looks at the consequences of the action, and asks whether they are good or bad. A lie may be right if it increases the greater good; or bad if it increases the greater harm.

Of course, one may combine different aspects of these theories: one need not adopt a pure virtue-based morality, but bring in utilitarian or duty based aspects, or a pure utilitarian morality, but incorporate aspects of virtue theory. Indeed, the Christian highly values the virtue of agape love, which is the active desire for the good, and leads one to do good for others. Thus, while Christian ethics is primarily virtue based, in practice it requires some consequentialist reasoning. However, we must adopt one of the three underlying methods as the first principle. Our premises are derived from that; the chains of logic may lead us elsewhere, but one of them comes first. Thus most moralists are primarily virtue, duty, or consequence based, and anything else is secondary.

Now, it does not necessarily follow that these three approaches are rivals. They may be different ways of reaching the same conclusion; indeed if they start from compatible premises one might expect that that would be so. The problem is that the adherents of each approach, and even different researchers in the same general school, do not start from compatible premises. Thus, ethics remains hotly contested; with each school quick to point out the flaws in the others (and usually less able to defend their own position).

In addition to these main approaches, there are some additional ways of studying ethics which have been used a great deal, but are on shakier ground. I will discuss a number of these, arguments from authority, human rights based approaches, contractual approaches to ethics and appeals to liberty or equality; before mentioning an important point due to Hume, and going on to outline some examples of virtue ethics, duty ethics and utilitarianism.

14.2.1. Arguments from Authority

> Is the pious loved by the gods because it is pious, or is it pious because it is loved by the gods?

So asked Socrates in Plato's *Euthyphro*. This has been used as the basis of a general argument against an authority based ethical system. The point is quite simple. Suppose that we have some authority G, who determines what is right and what is wrong. Then there are two possibilities:

(1) Either this determination is based on G's own arbitrary will (pious because it is loved by the gods), or
(2) It is based on some external standard (loved by the gods because it is pious).

(1) If it is based on G's arbitrary will, the ethical standard cannot be objective, nor is it open to reason. G may one day say, 'Do not murder', and the next 'Murder', and we will have to judge his utterances as well as we can. There would be no way to predict whether on a given day we should murder or not murder. This is not ethics. Furthermore, we must ask ourselves how we judge whether G is good or evil; what standard do we use? One based on G's own arbitrary prejudices? But this is meaningless; applying the standard to G gains us nothing. Something external to G? Then G is not the ultimate standard of right and wrong. Also, why should G be the source of our ethics, and not some other who has different views about committing murder? We have contradicted our original conjecture (14.1.4): there is no right and wrong because these descriptions are arbitrary and thus ultimately meaningless.
(2) If, however, G's pronouncements are based on an objective standard, then this standard is the ultimate source of ethics, not G. We have then contradicted our original assertion.

We have been left with a choice between ethics being arbitrary or objective, which either contradicts the rationality of ethics or our proposition that G is responsible for it. We therefore might well conclude that any ethics ultimately derived from authority is self-contradictory and fails.

This conclusion applies for almost any argument of ethics based on authority: kings, parliaments, clergymen, activists, lobbyists, treaties, declarations, constitutions, the rule of the mob, politicians or professors of philosophy; none of them have any right to dictate to us what is right and what is wrong based on their own whim or prejudice, because they each are a human authority. Each of them ought to be judged by an external, objective and rational standard if ethics is to have any meaning at all. They might, of course, deduce what that objective standard is from agreed premises and sound reason, and then teach it and apply it. But then they wouldn't be teaching on their own authority, or because they personally prefer that view. It is possible to accept something without liking it. Rather, they would be like a scientist teaching thermodynamics or special relativity, not just presenting his own subjective opinion to the students, but instead giving a (partial) description of objective reality.

Historically this Euthyphro dilemma has been applied to one authority in particular, God, which is somewhat ironic since God is one source of authority for which it does not apply. The reasons are obvious. First of all, God's lack of changeability means that the standard cannot change from one day to the next (and since God alone is unchangeable, this objection only applies to God), undermining the objection to the first horn of the dilemma. Since God is uniquely the universal creator and sustainer, there is a reason for singling Him out.

But the objection runs deeper than this. We have posed the dilemma as being a choice between the objective and the arbitrary; the first is seemingly fatal to God, the second seemingly fatal to reason. But we have to ask what we mean by 'objective.' Something is

objective and rational when its conclusion follows from an objective set of premises. What are the premises we can use to construct an ethical theory? There are two sources: firstly, the basic rules of logic, such as the principle of non-contradiction (which theists admit that God Himself cannot break, and thus thus this doesn't supersede the sovereignty of the theistic God); and secondly observations about human nature (what each man and woman have in common) and the nature of the universe. Now it is obvious that we must incorporate facts about ourselves into our ethics. There is no reason to suppose that the rules which govern beings who reproduce asexually will be the same as us who have two sexes, and any species which has three sexes or conceives communally will doubtless have other rules appropriate to it. Thus we must incorporate information about ourselves into our ethic. These are the premises from which an objective ethics must develop.

But why does the universe have the nature that it does, and why do we have the nature that we do? Because of God's arbitrary and, once selected, unchangeable choice. Thus objectivity is entirely compatible with God's sovereignty, and the arbitrariness of his choices. We can say that this possible world gives this ethics, and that possible world a different ethics, but God choose which of the worlds should be; and therefore selected what ethics should be. Thus ethics from the authority of God is both objective and His arbitrary choice, and because of God's relationship to the universe and his timelessness there is no contradiction between the two. Both reason and God's sovereignty remain. The dilemma is (in this case) false, because the two legs are not contradictory, and are combined in such a way that the pitfalls of both are avoided.

But this loophole obviously only applies to God. It may be there are other exceptions to 'ethics by authority'; I do not know of any. Men cannot define what is right and wrong by fiat; laws should be shaped to recognize objective morality, but cannot decree it.

Objection:
The above discussion stated that an objective ethics would nonetheless still depend on God. However, this is false as seen by the following argument.

If ethics is objective, then it is also independent of existence. For example, a moral statement might read, 'If such a being exists, then it ought to be like this' or 'If such a being exists, then it ought to act in this way'. This statement is true regardless of whether or not such a being exists. For example, the statement 'If there is a President George Hanover III of the United States, then he ought not to be assassinated' remains true even though there has been no President George Hanover III of the United States, and nor is there likely to be one. This means that ethics is independent of existence, and also, since the ethical statement proceeds logically without reference to God, also independent of God. Thus God is not necessary either directly or indirectly for a developed moral system. •

Response:
While it is true that a conditional ethical statement of the type described can be derived just from logical principles (assuming that morality can be objectively derived), nonetheless it is not the formulation of an ethical system. To use an analogy from the metaphysics of change in substances, it provides a potential moral statement, but to become an true ethical theory the potentiality needs to be actualised. For example, the statement 'If a Queldilope exists, then it also ought to be able to fulfil its final causes, which are ...' has no relevance unless there are such things as Queldilopes (and since I just made up the word, this is not the case). It remains a potential ethical statement. An ethical theory needs to be applicable, and such a statement is not so unless there exists at least one Queldilope. If such a being exists, then the statement is actualised, and becomes relevant. However, this actualisation ultimately depends on God. Therefore the objection fails because it does not distinguish between a potential or conditional moral statement and an actual or applicable moral statement. Every

actual moral statement requires that the being it describes exists, and thus this is not a reason to deny that ethics depends on existence, which would mean, since (as has been shown) all existence depends on God, ethics would depend (indirectly even if not directly) on God. o

14.2.2. Human rights

> While the Declaration of Rights was before the National Assembly some of its members remarked that if a declaration of rights were published it should be accompanied by a Declaration of Duties. The observation discovered a mind that reflected, and it only erred by not reflecting far enough. A Declaration of Rights is, by reciprocity, a Declaration of Duties also. Whatever is my right as a man is also the right of another; and it becomes my duty to guarantee as well as to possess.(Paine, 1791)

One of the most used, and most abused, systems used to guide moral thinking in the modern age is based on human rights (Veatch, 2007). We are told that all men are endued from birth with certain rights; and that to respect those rights is the duty of every member of modern society. We are reminded of the alleged rights to 'life, liberty and the pursuit of happiness'; of our supposed right to do what we will with our bodies; to marry the one we love regardless of who they are. Such views are sentimentally appealing, but we have to ask if they are grounded on reason. Where do such rights come from? Are they granted to us by the state? Inherent in human nature and recognised by the state? Primarily a restriction on the powers and authority of the state? How can we be sure that there is such a thing as human rights in the first place?

What I intend to argue in this section is firstly that the notion of human rights cannot provide the most fundamental principles of ethics. If it is to be useful, it has to be derived from some other principle. A corollary of this is that we cannot arbitrarily claim a particular right to whatever it is we want to have; that has to be proved from a deeper principle. Secondly, that if the notion of human rights is not foundational to ethics, then we ought to concentrate our discussion on that foundation, whatever it is. Everything that can be derived as a right can be deduced in other ways from that foundation. And that is how they should be argued for. The position I am attacking – that human rights discourse is the foundation of ethics – is certainly a minority one, but given the frequency of appeals to rights in contemporary discourse, it is nonetheless an idea I need to sweep away before moving on to more substantial ideologies.

The concept of rights derived from position, status or relationships – for example, the rights of Kings, the rights of the poor, the rights of slave-owners, the rights of parents, the rights of children – dates back to antiquity, and is there among our oldest written records. In each case, the rights were associated solely with the relationship between groups of people. The King had certain rights because he was a King; and those people who were not Kings lacked those rights. The subjects, also had rights derived from this relationship, rights which ought to be respected by the King (though historically frequently were not). The King had the right to collect taxes; the subjects the right to expect that these taxes would not be squandered but used as efficiently as possible for the general good of society. In essence, there was a contract (usually unwritten) between King and subjects, outlining the duties of each in that relationship, and the rights were derived from those duties and that contract. Of course, we can ask what the contact *ought* to be, or what a *good* contract would resemble, and these can be determined from the natural purposes of kingship: the ways in which mankind's natural tendency towards living in society and local communities is aided by having a King. Thus, in this sense, we can determine objectively a set of rights which Kings and subjects ought to possess on account of that relationship.

The concept of universal human rights, which people possess simply because they are people, is a far more modern idea. Our modern concept can trace it roots back to the late

middle ages, and the scholastic philosophers. Scholasticism had a well defined and thought out moral system - natural law theory - which was built on the concepts of formal and final causality (or what I call natural purposes). This system was developed without recourse to any notion of human rights. However, some of the later scholastic ethicists introduced the idea of human rights as a way of summarising the conclusions of this ethical system. To them, human rights were founded on the form and final causes of humanity; the concept is not fundamental, but was nonetheless seen as useful as an aid to conceptualise the results of the detailed logical arguments that build up natural law ethics. For example, the natural purpose of parents is to beget and raise children, and among the natural purposes of children is to be so educated and informed. This implies that children have the right to be brought up well by their parents, and parents have the corresponding right to respect and obedience from their children. The employee/employer relationship also leads to rights and responsibilities from both parties, ultimately derived from the contract between them which itself ought to reflect the societal natural purposes, which lead to the idea that every able bodied person ought to contribute to society at least what they consume from society. So employers have the rights from their employees of respect and to work productively; employees have the rights from their employers to a just salary, healthy working conditions, time to pursue their intellectual and family ends, and so on; together they have the responsibility to provide to society whatever good or service their company was set up to produce at a just and affordable price; and the wider society has the responsibility not to exploit or belittle their labour. All of these can be derived (though it is beyond the scope of this work to show how) from the natural purposes of the employer/employee relationship. Some of these rights and responsibilities are particular to people in that particular circumstance; others are found in many different circumstances, and thus we are drawn to the idea that there are universal human responsibilities, and from here it is a short step to universal human rights.

In the early modern period, the concepts of form and finality were abandoned, and with them natural law ethics (built upon them) had to go as well. This presented a problem, because some alternative basis of ethical theory had to be found. The concept of human rights was taken from late scholasticism, though stripped of its foundation. A new foundation needed to be supplied in its place. Probably the two most influential early modern thinkers in this regard were Hobbes and Locke. Hobbes viewed the moral code as a social contract; anarchic people coming together and mutually (albeit implicitly) agreeing to a way of living that allowed them to prosper in community better than they would have done on their own. The Hobbesian view of morality is thus something agreed on by the consensus of the people as a matter of convenience; human rights should be surrendered to this social contract. It should be clear that if this view of human rights is correct (and I believe that the continental idea of human rights as something bequeathed by the state on the people is based on it), then our moral code is based on very shaky grounds, being merely an argument from authority, and the worst authority at that, being simply the flights of fancy of the mob, and whatever rights we posses can be overruled by the vote of the politically powerful. It is entirely arbitrary what is declared a human right, and if arbitrary it is useless as a moral system. There is no objective way to resolve a conflict over what the rights are, except the rule that whoever shouts loudest (even if their shouting has the logical rigour of a leaky canoe) wins. Fundamental ethical principles should be based on better ground than this.

Locke, however, is a little more interesting. Locke's system of morality is based first and foremost around the idea of the ownership of property. His idea is that if you manufacture something yourself, then you have a certain ownership over that thing, and that ownership gives you the right to use it as you see best, and not as someone else would have it used. Ownership of property gives you certain rights over that property. Now we ourselves are the workmanship of God, and that gives God certain rights over us. However, God also gives us free will, so he in effect leases ourselves to ourselves. In this sense, we could be thought of as owners of our own bodies (though in reality our ownership is just a stewardship for

God); and this self-ownership gives us the property rights over our own body, and it restricts what other people can do to us or prevent us from doing. These property rights become Locke's version of human rights. The idea was immensely influential, particularly in the fledgling United States. The US declaration of independence, with its proclamation that it is self-evident that we are endued by our creator with certain human rights, and chief among these are the rights to life, liberty, and the pursuit of happiness, and the subsequent bill of rights are little more than a short step beyond where Locke finished. Of course, because we only lease our bodies from God, that restricts our own rights about what is acceptable for us to do. We cannot, according to Locke, use our self-ownership to justify immoral actions such as adultery or abortion.

But, of course, Locke's doctrine of self-ownership is in itself a controversial and dubious concept. What does it mean to own oneself? Ownership implies both an owner and the thing owned, and as these are two ends of a hierarchical relationship, there must be a distinction between them. Can we say that you (your mind) own your body, and make the distinction that way? At best that would only provide justification for bodily rights, and not the mental ones. But, worse, it relies on a Cartesian division between mind and body, which comes with its own difficult problems. So Locke's justification for the existence of rights fails, because it tries to prove one controversial idea by assuming a different controversial idea.

Locke's reasoning was rejected by most later philosophers, but his concepts of human rights and self-ownership were retained, only this time seen as self-evident rather than requiring a justification from more foundational principles. Indeed, with God thrown out by the embrace of first deism and then atheism, and formal and final causes rejected by all modern philosophers, it is difficult to see what other foundation human rights could rest on. It cannot be human nature or natural purposes, because the idea of a universal human nature on which to rest the universal human rights can only be a form of classical philosophy (or something that plays the same role), and this is forbidden to enlightenment philosophers (if we accept form, then we are back with Natural Law ethics if we accept finality or Plato's ethics if we don't); the other possibility is that it comes from a higher authority than mankind, which could only be God (either directly or indirectly through a regress of authorities), which atheists and deists cannot accept (and which leads to some form of divine command theory). Could human rights be accepted on practical grounds, because of their usefulness? No, because we would first have to define what is 'useful' without recourse to human rights. Whatever we use for this definition would form a more foundational moral principle than human rights, and it must ultimately lead back to either a universal human form or to God. And, of course, the concept of human rights might not be the only basis of achieving whatever practical benefit is offered. So an argument from utility cannot prove the existence of human rights to the exclusion of all other ideas. The only other possibility is that the concept of human rights is its own justification, that it is self-evident. But is it? That the idea has been absent in most of human history; that it would have been rejected by many of the greatest thinkers our race has produced; that even in Western society it first arose as a consequence rather than a foundation of moral reasoning, shows that it is not self-evident. The ideas that as a first moral principle we have inborn human rights, and own our bodies as property certainly aren't obvious to me, and I have never seen a valid justification of these claims.

So what is a right in the first place? It may, perhaps, be described as a principle of freedom or entitlement; a rule about what people are allowed to do or owed. The concept of human rights asserts that people have certain freedoms or entitlements as a consequence of being human beings. The concepts of humanity and these entitlements are interchangeable; there is something inherent in human nature which necessarily implies these entitlements. Note that this concept is implicitly restrictive; if there is a rule about what people are allowed to do, then it implies that there must also be rules about what people are not allowed to do; indeed it implies that the natural state of man is that he is not permitted

to do things unless he is given a rule permitting it. If we need to establish a principle of freedom, then that freedom is not inherent; for the inherent does not need to be established.

I should mention here the distinction between passive and active rights. A passive right limits other people's ability to interfere with your life. For example, freedom of speech is a passive right. It means that other people ought not prevent you from speaking. An active right is something which other people are forced to do something on your behalf. A right to an education, for example, can be interpreted either way. The passive right is that people shouldn't be able to prevent you from studying and learning the truth. An active right would be a right to a state education: you are forcing the state to do something on your behalf. Active and passive rights are different; and each requires their own defence.

It is clear that rights are distinct from abilities. We have the ability to commit murder, but few would grant us the right to commit murder. We have the ability to say what we will (and the natural tendency to do so), but why should we say that we have the right to say what we will? Now some would say that there is a difference between these two cases; after all murder harms another, while speaking freely at worst offends, and restricting such speech may harm the one trying to speak. But this objection presupposes an ethical standard (the principle of least harm) which must be proved correct before it can be used. Similarly every attempt to explain why some abilities are rights and others not must presuppose some previous principle, which cannot in itself be derived from the concept of human rights.

Another difficulty is that rights can only be granted or recognised by an authority. A mayor grants a man the freedom of a city; an instructor grants the learner the right to drive a car alone. A diploma grants someone the right to teach in a public school. We forbid this to others who have not passed the test or the exam. Yet if these rights are granted, then why not others, such as the freedom of speech or the right to marriage? Yet if they are granted by an authority, then an ethical argument based on human rights becomes an argument of authority, which we have already seen cannot be correct.

Another question we may raise is why speak of rights rather than responsibilities (or duties)? Now, it may be argued, as Thomas Paine did (Paine, 1791), that the two are equivalent: every right imposes a responsibility on others; every responsibility grants others a right. But if that is so, it is better to speak of responsibilities rather than rights. Rights are inherently selfish: 'I demand my rights'. To school someone in their own rights is to ask them to place themselves above others. On the other hand, to school someone in responsibilities teaches them to be selfless; to place the good of others above themselves. The latter inspires a much better attitude for life. Given that mankind is naturally selfish, we do not need encouragement in this regard; rarely do we meet someone who is too altruistic. Furthermore, to say that someone has rights is to say that their natural state is one of slavery, baring those few prescribed privileges (one is only granted the freedom of the city because most people do not possess such a freedom); to say that we are bound by responsibilities implies that we are otherwise free.

It is also not true to say that the language of rights and responsibilities lead to equal outcomes; two different ways of expressing the same thing. Take these two statements:

- We have the right to do what we will with our body.
- We have the responsibility to keep our body fit and healthy.

Now it is clear that these two principles cannot both be true, if applied in the same context as ethical postulates of the same weight[2]. For example, the glutton will attempt to eat himself to excess, damaging his body and his health. The person who argues from the 'right' stated above cannot object to this self-harm. The person who argues from the 'responsibility' stated above must object most strenuously. It is true that the first right can be rephrased

[2]Of course, one can avoid the contradiction by saying the first is a *legal* right and the second a *moral* responsibility; but my concern here is solely with the place of rights and responsibilities in ethical thought, so I only discuss them as moral rights and responsibilities. To determine what the law ought to be requires a sound ethical framework first of all; it is a question which can only be addressed much later in any argument.

as a responsibility (we must allow others to do what they will with their bodies); the second cannot be rephrased as a right, since the body considered in itself apart from the person cannot have a right, as rights can only be granted for rational agents. Now, if the two principles are in contradiction, we must choose one or the other, or neither of them. So we must ask where they are derived from.

Similarly, while every right can be expressed in terms of a responsibility, not every responsibility can be expressed in terms of rights (unless we extend the concept of rights to even inanimate objects). The most general principle should be used.

Equally, the concept of human rights presupposes a world-view that man has a certain inherent dignity; the concept of human responsibilities makes no such assumption (though does not contradict it). The approach which makes the fewest assumptions should be preferred.

Thus there are many reasons why, if we have to choose one of them, a language of human responsibilities should be preferred to a language of human rights.

But we may argue that rights concern civil law, rather than ethics. Not everything that is immoral ought to be proscribed by law. But civil law should be based on ethics. Law should only proscribe those things which are immoral, and should only encourage those things which are moral. Ethics should come first, then civil law. If the supposed right is opposed to ethical principle, then it should not be encouraged in law.

So an appeal to human rights is either an appeal to some human authority, which is useless, or an arbitrary assertion, which is even worse, or to some objective principle which is above human authority. But if the latter, then that principle should be used in the argument, rather than the appeal to human rights. In any case, the language of human rights should not contribute to any debate on ethics, expect at most as a summary of universally agreed principles. In particular, this language encourages intellectual laziness: every assertion of a human right is based on a fundamental ethical principle, which in turn is based on a fundamental metaphysical principle. Yet rarely do we see human rights activists do the donkey work of stating what that principle is, let alone establishing it and showing it to be reasonable. Perhaps they regard it as self-evident. It is not, however, self-evident to me that man possesses any inborn rights or dignity.

But the most important argument against the use of human rights language is that it entirely misses the point. The question of ethics should be about what is good; human rights completely bypasses this question; or rather it tries to present a resolution to this question without addressing the question. The question we should be asking is what it is good for a man to do (or what it is good for a man to be), and this is very different from the question of what do we have the right to do. The work of ethics should first be to decide what is good; and secondly what principles regulating action and character follow logically from this. Once we have done this, what need is there for the concept of human rights? Everything of substance is already accounted for; and if it is not accounted for, there is no rational basis for asserting it.

Now I do not wish to deny that human rights have not had an important historical role, and I fully agree with many of the principles in the declaration of rights, if not necessarily the reason why the principles are held. For example, there is the right to life (or that it is an evil to take an innocent life), the right to property (or that it is an evil to take someone else's property), the right to free speech (or that it is an evil to silence someone or to compel him to say what he does not believe), association (or that it is an evil to prevent people from meeting), and religion (or that it is an evil to compel religious or irreligious views on someone; or to hinder religious or irreligious practice – by all means try to persuade them, but not compel). All of these are crucial to a good society (albeit that they need to be balanced against each other). But one can reach the same place without using a doctrine of rights as a first premise. Saying that acknowledging human rights has a certain practical

utility is no defence that human rights actually exist. All the good that arises from the notion can be derived from other means, without mentioning rights.

Talk of human rights has one further use: it restricts governmental authority. Indeed, this is the chief purpose of such language. If we assert that all men have the right to a fair trial, then it prohibits the abuse of government power, so prevalent in tyrannies (indeed one of the first marks of tyranny) of fixing the courts to bias one side or the other. If we assert the right of freedom of speech, then it forbids governments from prohibiting what they wrongly see as hateful language (i.e. language that challenges their own, invariably at some level incorrect since none of us are infallible, world-view). But, again, we do not need the language of human rights to achieve this; a proper ethics of government and an immutable constitution limiting government inspired by this ethic has the same ability. Indeed, it is a better approach; for lists of human rights are policed by governments, and thus changed by governments, and therefore do not rule government but instead can be abused by a tyrant as a further tool of oppression against unfavoured groups.

What of the challenge that if we have no inherent right to speak freely, then we have no right to express the reasons for opposing human rights? There is a confusion here between cause and conclusion. The conclusion is the statement that people 'ought to be able to speak as they will', or its negative form, 'people ought not to be prevented from speaking as they will'. We can surely not deny this statement while making an argument against the concept of human rights. Now clearly, the statement 'People have an inherent entitlement to speak their mind', which is what the right to free speech asserts, leads to the conclusion. To argue from human rights not only implies the conclusion, but makes an additional statement concerning human nature; and it is this additional statement which those who oppose the concept of human rights object to rather than the conclusion that speech should be free. The conclusion does not imply the existence of this inherent entitlement.

That we can advocate for free speech while denying the right to free speech is clear, since there are other ways to argue for the freedom. For example, one may assert that people lack the right to prevent others from speaking. This is not an argument from entitlement, but rather from a lack of entitlement: the very opposite of the idea proposed by the advocates of human rights. The effect is the same as the argument from human rights, but it does not support the premises behind human rights thinking but rather undermines them. One may also argue for this conclusion from an ethical standpoint, that it is good for people to speak their mind therefore they ought to do so, either directly through a virtue based system, or indirectly through a utilitarian approach. This objection thus requires the assumption that only the premises of human rights thinking lead to the moral conclusion that one ought not be prevented from speaking freely; but this assumption is false.

For example, suppose we assert that we have a right to free speech. This is to implicitly imply a certain superiority in man; we have that right, and that right allows us to dictate the behaviour of others. But superiority implies that we are superior to someone, so who are we superior to? Well those constrained by the right, namely everyone else. But those people have the same rights as us, so simultaneously with our superiority to them, they are superior to us. This is not quite a formal contradiction, since one can be superior in some respects and inferior in others, but clearly to show logical consistency requires some effort. But suppose instead that we say that we lack the right to prevent others from speaking freely. Here the claim is that we are not superior to others (which is not the same thing as saying that we are inferior); we don't have the authority to restrict them. There is no chance of a contradiction here; it is merely a statement that as regard to speech, we are (all other things being equal) of equal worth. Thus a proclamation of a universal lack of human rights, while achieving the same end as a proclamation of universal rights, does so in a more consistent way, and a way more in tune with respect for the freedom of others.

In other words, saying that we have a right to free speech implies two basic principles. Firstly, that we in general ought to have permission to speak freely (as long as, for example,

we don't incite violence, and as the speaking is done in good order to ensure that everyone has chance to say their part at the most appropriate time – the sole exceptions I would make). Secondly that this permission is due to an inherent dignity within humanity. The first of these principles I will defend as vigorously as anyone. But the second principle I question.

The very idea that the rights of one person can be in conflict the rights of another undermines the whole concept of human rights as an ethical theory. Tensions such as this are invariably the sign of some fundamental problem in the idea; and in this case the problem lies in the whole concept of human rights. A true ethical theory should not have these tensions. Certainly, one can be forced to choose between two different goods, but this is different: it is not that both goods cannot in principle be held simultaneously, but only that due to limitations or the presence of some evil in practice only one of them can be pursued. The goods are not contradictory; only circumstance forces us to choose. But with these so-called rights, the contradiction is at a fundamental level, and can only be resolved by abandoning the whole structure.

To summarise: human rights are not a useful premise for ethical reasoning: each individual right is either an arbitrary assertion, or an argument from authority, or are based on objective more fundamental principles which should be stated instead. The only utility of a discussion of rights is as a teaching summary derived from some more fundamental ethical principle (as they were first introduced in medieval Europe). But even here, given that there is such confusion and misuse of the concept of rights, it is better that we instead fall back and rely on the fundamentals. The basic premise of modern human rights discourse, that we are born with a certain dignity which gives us these rights, needs to be demonstrated from first principles, and to my knowledge it has not been.

14.2.3. Contractual Ethics

Another approach, more commonly associated with laws but also sometimes to ethics, is to say that our legal or moral laws are determined by a social contract. People agree, explicitly or implicitly, freely a set of regulations which they live by for their mutual benefit.

There are several difficulties to the social contract view to ethics (and indeed to law, though I just wish to discuss the application to morality here). Firstly, there is the simple fact that it is not true in practice. How many of us signed a contract with society saying that we would subscribe to a particular moral view? Also do not each of us evolve our moral views as we increase our education or experience. Do we need to subscribe to a new contract whenever we change our mind? Or society changes its mind? Certainly I have no conscious memory of doing so.

Can people opt in and out of such a contract? Can, for example a murderer opt out of the clause in the contract which says 'Do not murder?' If so, then to murder would be moral, or at least morally neutral for him since there is nothing in his contract forbidding him (or, indeed, it would be morally acceptable to murder someone who failed to sign the contract). On the other hand, if the contract is imposed on the murderer by the wider society, then the contractual approach loses its justification that it was freely chosen by each member of society. The contract view either reduces to moral relativism, or an appeal to authority, whether the authority of the majority or the authority of those ancestors who first drew up the contract, and neither of these are tenable.

14.2.4. Appeals to freedom

A related view, pervasive in the modern world since the time of Hobbes and Locke (although the idea goes back further, perhaps to Ockham), is that freedom is among the most fundamental moral principles. A moral principle which appeals to human freedom (i.e. our ability to act as we choose) is instinctively seen as reasonable; a moral principle which seems to violate such freedom is considered dubious. The exception, of course, is when

one person's freedom conflicts with another, when some choice has to be made between the various liberties.

We must, of course, consider whether such appeals are truly based on solid ground, or at least under what conditions they are. There are, as we have seen, two aspects which define the capabilities of the mind: the intellect and the will. Freedom is (or seems to be in the modern understanding of the concept) primarily associated with the will. Objective moral standards, are, on the other hand, as a part of the truth, matters only accessible to the intellect. To assert the primacy of freedom is thus to assert the primacy of the will over the intellect. The alternative view, which is to say that human freedom is good as long as it is directed towards the good, states that the intellect ought to be primary, and the will guided by it. An appeal to freedom as a good in itself thus seems to me to be to be the height of irrationality - it is a side-lining of our rational faculties; a statement that irrational desires are more important than a knowledge of objective truth and its consequent application. That is not to say that 'freedom' in itself is a bad thing: on the contrary most of the time it is good; but it does mean that accepting as a first premise the positive 'I have my freedom' or the negative 'You are restricting my freedom' is not a valid form of ethical argumentation, unless it can first be proved as an objective truth that satisfying every impulse of the will is our guide to what is good.

But, against this, we see that ethics by definition leads to a set of rules saying which of our desires are good and which are evil, which carries the implication that some possible desires are evil. To assert that our desire is always good (unless it interferes with someone else's desire) undermines the whole study and purpose of ethics, and leads to a moral nihilism in which anything is permitted. Instead of accepting the principle of encouraging freedom as a premise, we must always go deeper, and ask whether it is good to permit or restrict that particular form of freedom. Most of the time it will, I am sure, be good to permit it; but we cannot say as a general principle that it would always be good.

Thus appeals to freedom, while emotionally seductive, are ultimately a means of distracting from and avoiding the real question we should be asking: 'Is this thing a good or an evil?'

Objection:
What? Are you some kind of apologist for tyranny? •

Response:
Now, it is certainly true that in Western Europe, North America and various other places around the world that the call to freedom and liberty was a key part of the overthrow of various corrupt governments and their replacement with better ones. And for this reason, I think, some people might associate a statement that the pursuit of freedom is not a good in and of itself with an attempt to defend the extreme actions of some dictator. We should, however, be wary of such arguments. That the principle had good results in some circumstances does not mean that the principle is good universally. There are other reasons to oppose tyrannical government.

Secondly, I am not arguing about what ought and ought not to be illegal, but what is and is not immoral. Not everything that is immoral ought to be illegal. To say that certain desires are bad does not mean that a government ought to restrict by law the actions associated with those desires. Nor does it say the opposite; it remains an open question. However, a tyrannical government acts by means of the law. It may be immoral for a government to pass a law proscribing something which is itself immoral.

Thirdly, the primary purposes of government are to provide various services and utilities which cannot be efficiently supplied through private means; to respect and enforce private contracts and provide protection against fraud; and to protect its citizens and to allow them to easily go about their business in pursuit of the various goods, both the general goods

relating to all mankind, and those specific goods related to their position (for example, their job) in society. A good government is one that fulfils this purpose. Thus a good government is one that serves its citizens and makes it easier for them to run their businesses, become prosperous, raise families, or otherwise seek out goodness for themselves and others. A tyrannical government, on the other hand, seeks to dominate its citizens and seeks the enrichment of its own members as a greater priority than the enrichment of the common man. Thus a tyrannical government is, almost by definition, an evil. In this way, the moral principles which I am developing argue against tyranny, and not in favour of it. I just have a different reason for opposing it than was used historically.

Fourthly, and finally, the enthronement of freedom above goodness has not always led to good governance. Not every tyrant takes the form of a French King or British tax collector. For example, as I write, most governments in Europe and North America enthusiastically support the notion of personal freedom, but nonetheless pass numerous petty regulations which stifle business and hinder those private men trying to do good in society, while simultaneously enriching civil servants and politicians at the expense of the taxpayer through (amongst other things) corruption, centralisation, inflated wages, and the creation of well paid but pointless jobs. Thus, the raising of freedom as a primary moral principle only inhibits some forms of bad governance, but not all of them; while a proper moral argument will target all forms of bad governance. However, in as much as the argument from freedom distracts from the more sound argumentation which can also defeat these types of tyranny as well as the more despotic, it is the elevation of freedom that is more supportive of the tyrants than a more rigorous approach to ethics. ○

14.2.5. Appeals to Equality

Another approach taken to ethics is to appeal to the principle of equality, which states that 'Like should be treated alike.' This principle is good, and I do not know of anyone, whether on the political left or the political right, who objects to it. There is, of course, a corollary, which is less frequently appealed to, which states 'What is unlike need not be treated alike,' and many would go further and say 'What is unlike ought not to be treated alike (to the extent to which it is truly unlike).' However, if we accept the corollary, then I think we must also accept its extension, otherwise our principle of equality would read 'Everything should be treated alike', which nobody accepts. For example, we do not treat a serial killer in the same way as we treat a charity worker (all other things being equal), and nor should we. We don't go to a brain surgeon to marvel at their paintings, nor to an artist if we need a tumour removed from our heads.

I want to highlight three problems with using the principle of equality as a guide to ethics. The first is 'How do we determine what is like and what is unlike?' Let me offer an example of what I mean: selective education. The idea behind this is that at a young age (say 11), children should take an exam, and those that pass it should go on to a grammar school, where they will be granted an ideal learning environment, while those who fail it are sent to a secondary modern school, where the speed of learning will be slower. Now let's look at this from the principle of equality. One perspective states that this is unequal, because you are treating some children differently from others. Some children receive a good education, while others a bad education, but ultimately they are all just children, all just the same. It is thus unjust to separate them. However, if we were to put them all into the same school, then all the children would learn together, at the same rate, and it would be equal. The other perspective states that, on the contrary, the grammar school system is the one that fulfils the principle of equality, because every child learns at the rate best suited to their own abilities. The grammar school students learn at the faster speed, but they can cope with it. The secondary modern students learn at a slower pace, but again, they need to, because they are not so quick at grasping new information. A comprehensive

system, on the other hand, is unequal, because here the pace that the teacher goes at will be some compromise, and the brighter students will not be sufficiently stretched, while the less able students will be stretched too much. Thus the students are treated differently, and the principle of equality is violated. Of course, this is just an example: there are far better reasons for and against selective education, which I don't want to go into here. But does the principle of equality mean that every student should be educated at the pace which is best for them, or that every student should learn at the same place? I don't think there is an easy answer to this, because the children are like in some respects and unlike in others.

A similar issue arises in, for example, homosexual marriage. Proponents of this say that the principle of equality supports them, because they treat same sex attracted people in the same way as an opposite sex attracted people. Opponents (though this is not their strongest argument) say that homosexual marriage violates the principle of equality, because it treats two different types of relationship as though they were the same, bringing opposite sex relationships down to the level of same sex relationships, effectively stripping the opposite sex relationship's most important feature, the ability to naturally procreate, from the definition of marriage. The difference is not over the principle of equality, but over whether the two circumstances are truly alike in all respects that matter to marriage. If they are, then same sex marriage is 'equal marriage'; if not, then it is in reality 'unequal marriage': same sex marriages will always be second class because they lack an essential purpose of marriage, the ability to naturally procreate. And, of course, denying homosexual marriage need not be seen as denying the principle of equality, if we accept that men and women are different. Every man is treated equally: able to marry a woman, but not permitted to marry a man. Since men are not like women (in this respect) there is no violation of the principle in saying that one can be married and the other not. It is only by adding to the principle of equality the idea that men and women are alike with respect to their roles in marriage that someone can conclude that equal marriage implies homosexual marriage. It is on this second principle that the two sides of the argument disagree; so that is what the debates and slogans should focus on. Of course, the two sides in the debate also have different understandings of what marriage is; they both view the other as either immoral or idiots or more usually both, because they each accept very different premises.

Thus before we can judge whether two people or two groups are alike or unlike with respect to a particular institution, we need to know what that institution requires; which means that we need to know what its purpose is and how it ought to be fulfilled. Thus prior to applying the principle of equality, we need a well thought out set of definitions, including the definition of what is good with respect to the institution we are considering. But those definitions will be sufficient to derive a complete ethical theory, making the principle of equality redundant.

A second problem is that treating like people alike is not necessarily good. What if we starved everyone in the world? We would be treating them alike, but we would not be treating them well. Thus that we are treating people equally is no guide to whether or not we are promoting good.

Thirdly, and related, is that the principle of equality doesn't tell us how to treat people. We should treat two schoolboys the same. Fine. But should we give them both schoolbooks, or venomous serpents, or both of these, or neither? The principle of equality by itself doesn't tell us, and nor can it. It therefore cannot be the most fundamental ethical principle (though it might be something derived from that principle).

14.2.6. Voluntarism

Voluntarism tries to overcome the problem of ethics by making human desire the guide for what is good. It represents the superiority of the will over the intellect. One form of this is that we should, as much as possible, let people decide for themselves what is good for them, and let them choose their identity for themselves. As with all ethical fallacies,

this takes the core of a good idea, and exaggerates it beyond all sense; the problem is not so much in what this philosophy affirms, but in what it denies. The alcoholic wants more drink: should we give it to him, knowing the harm that it does to both himself and others? A pure voluntarism says that we should. The problem is that if we try to qualify voluntarism, either by saying that at a later time he would appreciate our actions in withholding the drink, or if he were in his right mind he would not ask us for this, then we are basically admitting that the person's will is not the highest principle. How do we know what he would want or not want in the future? How do we know what constitutes his right mind? We are effectively transferring the moral standard from his own will to our own will. But then, how do we know that our own will is good? How do we judge between the desires of one person's will and another when they clash?

And this is the fundamental problem with a voluntaristic ethics. Ethics is there to judge which of our actions, desires and character traits are good or evil. Voluntarism defines what is good and what is evil by using our actions, desires or character traits. It is a wholly circular position; it cannot answer the question: is this desire or attraction I have good? Is this desire or attraction that that man over there has good? Or rather, the only answer it can give to such questions is 'yes'. And this basically destroys the whole study of ethics, because it means that there can be no general moral rules. A voluntaristic society can only be wholly irrational.

Our moral judgements will, of course, be made based on both our will and intellect. But when these come into conflict, and cannot compromise, one of them has to be given priority. The intellect can reach out to objective truth, and is firm. The will is constantly changing and subjective. If we are to develop a sound ethical principle, the intellect has to be given priority over the will. Someone's identity is not what they define themselves to be or desire themselves to be, but what they objectively are. And if their desires or self-definitions clash with that, then it is their desires, not the rest of them, that represent the problem.

14.3. Is an Objective Ethical Theory Possible?

14.3.1. Can We Construct an Ought from What Something Is?

One of the most dominant thoughts in the modern understanding of ethics was first stated by Hume:

> I cannot forbear adding to these reasonings an observation, which may, perhaps, be found of some importance. In every system of morality, which I have hitherto met with, I have always remarked, that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when of a sudden I am surprised to find, that instead of the usual copulations of propositions, is, and is not, I meet with no proposition that is not connected with an ought, or an ought not. This change is imperceptible; but is, however, of the last consequence. For as this ought, or ought not, expresses some new relation or affirmation, it is necessary that it should be observed and explained; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it. But as authors do not commonly use this precaution, I shall presume to recommend it to the readers; and am persuaded, that this small attention would subvert all the vulgar systems of morality, and let us see, that the distinction of vice and virtue is not founded merely on the relations of objects, nor is perceived by reason (Hume, 1739).

This passage expresses the conclusion of a section on whether morality can be derived by reason. Hume argues that it cannot be. I will refer to it as the *is/ought* problem; the

idea is also known as the *fact/value* dichotomy. The basic premise of this argument is that describing what it is to do something bears no relationship with what we ought to do. For example, we can split a description of giving food to the poor into its constituent parts, understand the mechanics of those parts in full detail, but in none of that process is there any reference to moral value; nor would a similar argument if we were to take food away from the poor contain any more than a description of the situation. To get a moral conclusion that one is good and the other an evil we need to add something beyond the mechanics of the situation which we can induce from observation. But Hume is an empiricist: he believes that aside from basic logic and mathematics, we cannot know anything beyond that which we directly observe. Therefore the gap between empirical fact and moral value is unbreachable. We cannot tell from observation, experience, or the orientation of our desires whether an action is good or evil. Thus far, I fully agree with Hume, and cannot see how his conclusion can be avoided; his argument destroys most modern approaches to ethics. The question is whether there is a way to determine value which Hume did not consider, a definition of moral principle where his argument is inapplicable.

Now, it should be obvious that Hume's argument is based on various premises. The first is Hume's empiricism, although obviously other forms of this argument exist in many non-empiricist philosophies. A simple empiricist argument for the conclusion that there is no objective ethics is that: a) An objective ethical standard is only possible through comparison against some ideal; b) such an ideal is not directly observable, or at least it cannot be directly observable that it is an ideal; c) Neither is it derivable from basic mathematical, logical or geometrical axioms; d) the assumption of empiricism, namely that only that which is directly observable or logical, geometrical or mathematical truths can be known; and the conclusion e) that no objective ethical standard can be known is claimed to follow. Hume's argument is more sophisticated than this (and such an argument needs to be; for example we also need to exclude that a moral ideal can be found through a combination of observation and mathematical reasoning; and to consider different types of so combining them beyond the mere induction and deduction that Hume understood; and it would also need to better parametrise degrees of uncertainty in knowledge), but it relies on empiricism in the same way. Of course, as discussed earlier, empiricism is untenable: if we accept empiricism, we have to reject theoretical physics, history, and many other fields which form knowledge based on both observation and reasoned argument. If we bend empiricism enough to permit this sort of theoretical knowledge, then we also bend it enough to destroy this argument against an objective ethical system. Thus this dependence of Hume's argument on his wider philosophy is a considerable weakness for his version of the argument.

Hume's second premise is that he adopts either a duty based or consequentialist ethic: he is more concerned with what we ought to do rather than what we ought to be; that last question he neglects entirely. But again, his argument can be adapted to a virtue based system if his remaining premise holds.

That premise is that we can break down an act into the individual microscopic elements, and those elements do not inherently contain any reference to value or goodness. Now, in an enlightenment philosophy based on mechanism or something similar to mechanism, I believe that this is correct. Hume's is/ought dichotomy holds, and the moral philosophers of the enlightenment have a serious problem: if we cannot reason from objective observation to moral value, and the basic truths of logic and mathematics are insufficient to reach a moral system, then there is no objective source of morality. We are, in effect, reduced to either an argument from authority (of which only God's authority is legitimate, but that is rejected by secular philosophy), or to relativism or voluntarism, which, if consistently applied, mean that there are no binding moral rules at all since everyone can decide which rules they themselves want to follow. The one way to circumvent this conclusion is to say that moral rules can be derived solely from logic and mathematics, accepting as axioms only

the principle of non-contradiction and various other self-evident logical principles: this is what Kant attempted, but his effort, though worthy, is generally regarded to have failed.

However, Hume's acceptance of the enlightenment principle that there is no formal or final causality is crucial to his argument. In Aristotelian terminology, form relates to a description of what something is, while finality relates to the effects that would normally proceed unless hindered (for example, as a tree has the final cause of bearing a particular type of fruit). This hindrance might be external to the being (for example, if we stop a tree from bearing fruit by chopping off its branches), or it might be due to a defect in its nature (a particular tree happens to be seeded infertile, or made infertile through some disease). The Aristotelian definition of goodness is that there are no such hindrances to the final cause being achieved; each hindrance is a particular evil. Thus goodness, or value in the terminology used in modern philosophy, is related to final causality. What something is, the fact of its existence, can be found by comparison against the form. If we accept this particular framework, then Hume's premise entails the denial one of the two remaining steps of the argument: either we must deny any link between formal and final causality, or we can deny that we ought to encourage goodness (as defined by Aristotle) in ourselves and others.

The first of these steps, however, is a key aspect of Aristotelian metaphysics. In my own version of the metaphysics, the form is determined by the energy bands of the appropriate effective Hamiltonian, that Hamiltonian is written in terms of various creation and annihilation operators, and those creation and annihilation operators give the possible final causes for each state, and because those final causes are not symmetric (i.e. the probability of moving from A to B is greater than that of moving back from B to A) we can say that there is a natural direction in the development of the organism, and thus (after a complex calculation) from the effective Hamiltonian of the tree we both derive its form and conclude that trees have the natural tendency to bear fruit. One can start from the form, construct the corresponding Hamiltonian, and from that derive the final causes, or start from the final causes, and work one's way back to the form. Thus there is no gap between form and finality which Hume can use.

What about the second assumption, that we ought to encourage goodness? In the Aristotelian framework, to be good is to be a manifestation of what one truly is. For example, we can start by specifying a certain set of purposes which underlie what it is for a being to be good, derive from that the Hamiltonian of a being which satisfies those purposes, use the principle of continuity to construct the form of that type of being. Some of those states will more easily achieve those purposes than others. The principle of encouraging goodness means that we should encourage the being to be in a state which more readily achieves its final causes. But what is the alternative to encouraging goodness? We can either actively discourage it, or be ambivalent about it. To actively discourage goodness is to encourage it to be able to fulfil a different set of purposes; either mould it to some other purpose, or to suppress one of its natural purposes. But an expression of the natural purposes was the starting point of our construction of what that being is. To actively suppress those purposes is to try to turn the object into something else. Thus to discourage goodness is to encourage something to be something other than what it is. But that means that we are fighting against the principle of non-contradiction. What about if we are simply apathetic? Here we are merely content for something to be that which it is not; the apathy is better than the actual discouragement in the sense that to be content with being irrational is better than to strive to become irrational; but neither of them are ideal. We thus conclude that it is irrational to do anything but encourage goodness. We are, however, rational creatures, and thus to seek irrationality is to deny our own nature. Thus if we are to be what we are as rational creatures, then we ought to encourage goodness. But the first part of that clause is a tautology, so it can safely be omitted. We ought to encourage goodness, as defined by the Aristotelian framework.

Thus the denial of formal and final causes, and the link between them, is the key premise on which Hume's argument depends. Deny form and finality, and there is no ethics (except possibly divine command or some Kantian system, if a complete Kantian system can be constructed); accept form and finality and we immediately have the basis of an objective ethical system. But we have seen that acceptance of form and finality implies the existence of God, and I will show that Kant's system fails. Thus every secular system of ethics finds Hume's *is/ought* dilemma an insurmountable barrier; to the two religious ethical systems, divine command and natural law - which each reject key premises of Hume's argument - it is no barrier at all.

14.3.2. Moore's naturalistic fallacy

Hume's thesis, that what is moral cannot be derived from what is reasonable, was taken up in a slightly different form by G.E. Moore (Moore, 1903), who labelled it as the *naturalistic fallacy.*

> Yet a mistake of this simple kind has commonly been made about good. It may be true that all things which are good are also something else, just as it is true that all things which are yellow produce a certain kind of vibration in the light. And it is a fact, that Ethics aims at discovering what are those other properties belonging to all things which are good. But far too many philosophers have thought that when they named those other properties they were actually defining good; that these properties, in fact, were simply not other, but absolutely and entirely the same with goodness. This view I propose to call the naturalistic fallacy and of it I shall now endeavour to dispose (Moore, 1903, § 10).

This is used, for example, to assert that one cannot derive what is good from what is pleasurable or what is desired. I fully agree with him about this. But Moore takes it further, and in a direction with which I strongly disagree, to contend that the concept of goodness is ultimately undefinable.

> (1) The hypothesis that disagreement about the meaning of good is disagreement with regard to the correct analysis of a given whole, may be most plainly seen to be incorrect by consideration of the fact that, whatever definition may be offered, it may always, be asked, with significance, of the complex so defined, whether it is itself good. To take, for instance, one of the more plausible, because one of the more complicated of such proposed definitions, it may easily be thought, at first sight, that to be good may mean to be that which we desire to desire. Thus if we apply this definition to a particular instance and say When we think that A is good, we are thinking that A is one of the things which we desire to desire, our proposition may seem quite plausible. But, if we carry the investigation further, and ask ourselves 'Is it good to desire to desire A?' it is apparent, on a little reflection, that this question is itself as intelligible, as the original question, 'Is A good?' - that we are, in fact, now asking for exactly the same information about the desire to desire A, for which we formerly asked with regard to A itself. But it is also apparent that the meaning of this second question cannot be correctly analysed into 'Is the desire to desire A one of the things which we desire to desire?': we have not before our minds anything so complicated as the question 'Do we desire to desire to desire to desire A?' Moreover any one can easily convince himself by inspection that the predicate of this proposition – good – is positively different from notion of desiring to desire which enters into its subject: 'That we should desire to desire A is good' is not merely equivalent to 'That A should be good is good'. It may indeed be

> true that what we desire to desire is always good; perhaps, even the converse may be true: but it is very doubtful whether this is the case, and the mere fact that we understand very well what is meant by doubting it, shows clearly that we have two different notions before our mind (Moore, 1903, § 13).

He continues that to say what is natural, as in what is normal in nature, cannot be regarded as the uniquely defined standard for goodness; the place where evolution leads us cannot be regarded as that standard. He dismisses Metaphysics as the grounds for ethics. Firstly he distinguishes the questions of 'What ought to be?' from 'What ought we do?' and claims that Metaphysics can influence answers to the second question by, for example, positing a life after death, and what sort of actions we conduct now might influence our place in the next world.

However, for the most part, he believes that cases such as this are the exception, and that it is clear that the usual problems pondered by metaphysicists bear no real relation to practical ethics, since they tell us nothing of future reality but only of eternal principles. There are two aspects to practical ethics: being able to predict the future, and discerning which future situations are good. Consider first the role of metaphysics in predicting the future. The metaphysicist states that nothing is real except the eternal. In particular, the Metaphysical principle that only that which is eternal is real and that our actions only affect the future implies that our actions have no moral consequences since the future is not real. For the metaphysics true eternal reality to impact our actions it thus cannot be the only true reality. 'For a moral rule, bidding us realise a certain end, can only be justified, if it is possible that that end should, at least partially, be realised. Unless our efforts can effect the real existence of some good, however little, we certainly have no reason for making them. And if the eternal reality is the sole reality, then nothing good can possibly exist in time: we can only be told to try to bring into existence something which we know beforehand cannot possibly exist.' (Moore, 1903, § 68) Secondly, he argues that the metaphysical eternal reality cannot be the sole good, since its realisation in which a good is achieved must also be a good reality. 'That the eternal reality is good, will by no means justify us in aiming at its manifestation, unless that manifestation itself be also good.' Thus the eternal principles which Metaphysics describes either cannot be the exclusive set of goodness – it cannot contain the whole of possible goods – or no goodness can be realised. Therefore metaphysics can't really add anything to this question of what future consequences our actions will have.

Moore's argument here is flawed in several respects. Firstly, he assumes a consequentialist view of practical ethics that a goodness of an action depends on the future consequences of that action. Thus this argument has no effect on any duty or virtue ethics derived from metaphysical principles. Secondly, he only considers metaphysical models which posit some eternal reality which possibly has some manifestation in the physical world: perhaps he has in mind Plato's model of forms, or something similar to it. However, there are other metaphysical models which have different approaches; for example the scholastic view of substance as the union between the abstract form and matter. Every material object has an abstract component. In this view, making a difference between the eternal world of metaphysical abstractions and the physical world of real objects is a false dichotomy, since the abstract idea and physical object are necessarily united. As Moore admits, this whole argument is a straw man anyway, since those who advocate metaphysics as a guide to what ought to be done concentrate on the other issue – What is good? Predicting the future is the domain of physics; determining goodness where metaphysics comes in.

So what of this second point? Can metaphysics tell us what is good? He answers 'No'. Why he answers this is less clear to me. He appeals to his earlier arguments, and his coining of the naturalistic fallacy, but these did not discuss metaphysical approaches of defining goodness. 'If we ask: What bearing can Metaphysics have upon the question, What is good? the only possible answer is: Obviously and absolutely none.' (Moore, 1903, § 69) But

for this assertion, Moore offers little evidence. He ponders about the difference between the questions 'What existing things are good' and 'What sort of things are good', asserting that the first question reduces to the second and the question 'What is real?' The second question is aided by the confusion between 'This sort of thing is good', or 'This thing, if it existed is good', or 'This existent thing is good.' The same words are used when distinguishing between something which exists and which might exist. This ambiguity, Moore asserts, is the source of the problem: metaphysics is the study of being as being, so deals with questions of what is real. So to the question 'This existent thing is good', it adds in some part of metaphysics related to existence; and metaphysicists then use this to claim that they also have something to say regarding goodness. Metaphysics only refers to eternal truths, but saying 'This thing as an eternal abstraction is good' cannot be the same as saying 'This particular thing is good', since the goodness of particular things depend on the surrounding circumstances.

> When, therefore, we are told that the future realisation of the true self is good, this can at most only mean that the future realisation of a self exactly like the self, which is true and exists eternally, is good. If this fact were clearly stated, instead of consistently ignored, by those who advocate the view that the Supreme Good can be defined in these metaphysical terms, it seems probable that the view that a knowledge of reality is necessary to a knowledge of the Supreme Good would lose part of its plausibility. That that at which we ought to aim cannot possibly be that which is eternally real, even if it be exactly like it; and that the eternal reality cannot possibly be the sole good – these two propositions seem sensibly to diminish the probability that Ethics must be based on Metaphysics. It is not very plausible to maintain that because one thing is real, therefore something like it, which is not real, would be good (Moore, 1903, § 70).
>
> It is now plain that a metaphysical principle of Ethics which says 'This eternal reality is the Supreme Good' can only mean 'Something like this eternal reality would be the Supreme Good.' (Moore, 1903, § 70)

This would only be useful if the future were some imaginary utopia. But in reality, this eternal standard would never be realised – it is in essence a fiction. Metaphysicists at best would describe some ideal world that has no bearing on the real world, and can't be compared to it. Thus, Moore concludes, metaphysics cannot have a role in ethics.

Once again, Moore's argument is very limited; he basically presupposes a vaguely Kantian separation between the abstract ideal of metaphysics and the real world, and thus inherits Kant's flaws. He does not consider that our abstract notions might cover a range of possibilities, so that reality can coincide somewhere within that range. Of the scholastic conception of form, that there are numerous different ways a being can be actualised, he has no concept. The question is not 'What sort of things are good?' But 'Which of the various possible states are good for this particular type of thing?' Because the ideal is not a point in the space of beings but a region covering numerous different states, there is no reason why that ideal may not be realised in the real world; or that we may not measure the distance from that region to the actual state. More importantly, Moore makes no mention of the key metaphysical concept that links form with goodness – that of final causes, nor the relationship between the two. In short, Moore presumes that there are no formal nor final causes. He does not consider that form might refer to a set of potentia, and he has no consideration of the role of final causality. His arguments are based on presumptions (such as consequentialism, a division between the abstract and the material) which scholastics would reject. For metaphysical systems which accept the enlightenment presumptions, Moore might well have a case that goodness is undefinable. But he does not respond to the Aristotelian definition of goodness other than to assert that the undefinability of goodness

follows from a set of premises which Aristotle's disciples would deny. Moore's discussion thus presupposes only certain particular metaphysical views; and has no validity against those, for example, which deny his contention that metaphysics only concerns some unreachable eternal reality.

Moore's thesis does not on the face of it seem to be related to the is/ought dilemma; it is simply the thesis that a) the concept of goodness is the ground of ethic thought; and b) that good is in itself indefinable. The two can be connected if we presuppose (as, I think, Moore does) that after defining what is a good object we are then given an understanding of good actions; and then that we ought to do is what is good. If goodness is indefinable, then it cannot follow from some basic principle related to empirical or tautological fact, for this would give it a definition. So for Moore, *ought* can only be determined from *good*, but a sensible definition of *good* cannot be derived from *is*.

Of course, Moore's argument has numerous problems and difficulties. The simplest (and perhaps most naive) is this: Moore argues that the concept *good* is undefinable. So did he make a *good* argument?

For if we answer 'no' to this question (or not sure, or that the question is meaningless, which amount to the same thing), then clearly we have no reason (from this argument alone) to jettison the concept of goodness. On the other hand, if we answer 'yes', then either our answer 'yes' is meaningless because it is based on a meaningless or subjective definition of what constitutes a good argument (in which case again we again have no reason to accept Moore's conclusion: since this conclusion is intended to be objective, applying in every circumstance, and not subjective), or we have in our affirmation of it undermined Moore's conclusion.

The obvious criticism of this naive objection is that moral goodness and the sense in which an argument is good are two different things; the same word is used to describe different concepts. Thus we can argue that goodness can be defined in certain senses, but not morally. But then why make this distinction? Why not use the same sense of goodness in judging an argument (where the concept is well defined) as we do when we judge a person? A good argument is one that fulfils its purposes (namely to convince people of the truth of a particular viewpoint), that is one that proceeds from sound premises through solid reasoning to reach a truthful conclusion, and expressed clearly enough that it can persuade others. Why should we not adapt the same concept to people - that a good person is one who fulfils his purposes? The only difference is that it might be harder to identify what those purposes are. But that it requires extra work does not mean that it is an altogether different concept; the difference is only of degree rather than kind. I think that we can only deny that this sense of goodness can be applied to people if we suppose that it is impossible that there are natural purposes for people. This would only appeal to those that erroneously deny final causality. For if we accept final causality, then the purposes[3] can ultimately be derived by considering the final causes (which are ultimately what we use to define person-hood) – we only have to find these purposes (or a subset of them) to have a solid and objective grounding of ethics. Thus this idea – that the word 'good' cannot have the same sense when applied to (for example) an argument and a person – relies on the premise that there are no final causes for natural objects. It is only reasonable for one who has bought into the underlying presuppositions of the enlightenment. Yet, as I have already shown, those premises are unsustainable in the light of modern physics. Thus this criticism of my naive objection fails. We can, in principle at least, use the same sense of goodness with respect to an argument as we do for the moral worth of a person. And thus Moore's whole thesis is obviously self-refuting.

[3]Using the word 'purposes' here is a little misleading, since some might think that it implies an external will, contradicting my previous definition of final causality. I shall define precisely what I mean by *natural purpose* later, but, like final causality, it still maintains the sense of a natural tendency towards some end.

14.3.3. The unnaturalistic fallacy

If ethics is not a matter of reason, then it must be, by nature, an issue of subjective impression; some sentiment only.

Hume is left with regarding ethics as being derived from pleasure and pain; ultimately on whether we regard things with favour or disfavour. Now it is not self-evident that virtue and vice are inseparable from pleasure and pain. What is pleasurable is often evil. Taking drugs is (so I am told) pleasurable, yet it is a great evil, as it damages the mind and distorts the appetites, making us addicted to and our will focused on something that is ultimately worthless. Unless we have a well-defined understanding of good and evil, we cannot make this assertion.

It is also clear that an ethics based on judgement cannot be eternal or universal. It will vary according to the passions of each age or culture, as these may favour different pleasures or despise different pains. We would not be able say that to hate homosexuals is universally bad, if ethics were based on subjective judgement over matters of pleasure and pain (a view I reject although it is accepted by some[4]). If a culture gains pleasure out of hating (as some have and some still do), and this pleasure outweighs the pains of the hated, then how can those who believe that goodness is merely an impression condemn this impression? They would either asserting that their own arbitrary standard is better than someone else's arbitrary standard, which is nonsensical; or say that there is an objective standard above each group and culture's own impressions, but then this standard must be based on something other than impression, and reason, denied to us by our primary premise that *ought* cannot be derived from *is*, is the only other standard available.

Ethics is intended to provide us with an external standard against which we can judge our appetites. Is our enjoyment of a play or novel good or bad? Is our enjoyment of the company of those we love good or bad? To say that it is good because we enjoy it and we enjoy it because it is good is circular reasoning. We cannot judge whether the enjoyment is good without a standard of goodness external to that enjoyment. The alternative is, then, that goodness is based ultimately on human judgement; but in this case it is meaningless. Ethics cannot regulate human desires or passions if it is built on those same human desires or passions. It is fallacy to say that we can derive a meaningful description of what *ought* to be from anything except that which *is*. I shall label this as the unnaturalistic fallacy. If we can't in some way link *is* to *ought*, then ethics is nothing but a mirage, and all morality abandoned, and to my mind there are few things more certain than that.

Now, this statement does not contradict Hume's that *ought* cannot be derived from *is*. But if they are both true, then ethics is meaningless; there is no right or wrong; we cannot criticise people for immorality because there is no morality. We are left with Nietzschean anti-ethics. Of course, I do not accept this unpalatable situation; since I utterly reject Hume's contention.

14.3.4. Summary

Philosophers often claim that Hume and Moore between them showed that there is no objective ethical standard. This claim is false. What they showed was that there is no objective ethical standard consistent with their premises and assumptions. So what are these premises? The first is that we need information concerning the nature of things as premises in our ethical theory. The second is the enlightenment presumption that the nature of things contains no inherent directionality or purpose. The first of these premises can be denied either by supposing that the axioms of logic alone are enough to derive an ethical system – which leads to Kant's system of ethics – or by supposing the axioms of ethics can be provided by some authority - and since only the authority of God survives the

[4]I would also disagree strongly with the notion that hating homosexuals can be anything other than an evil of the most degenerate degree, although the reasons why I believe that and what such a belief entails in practice will have to wait until I have established the foundations.

Euthyphro dilemma this leaves us with just divine command ethics. The second premise can be denied by supposing that there is some directionality or purpose inherent to the nature of beings, which leads us to the natural law ethics of the medieval scholastics (inspired by Aristotle). Given that Moore also responded to Plato's system of ethics, what Hume, Moore and Socrates (in response to Euthyphro) actually showed was that the only tenable systems of ethics in principle are those of Kant, or some form of divine command or natural law ethics.

14.3.5. Can We Construct An Is From An Ought?

In assessing Hume's dilemma, we must start with the question of what do we mean by *ought*. A moral statement is given by a simple statement concerning the word *ought*; so, for example, a virtue ethicist might say 'Mr Robinson ought to be courageous;' while a duty or utilitarian ethicist might say 'Mr Robinson ought to go out to fight against the pillagers'. The question Hume is raising is how we can justify this ought based on facts. So 'Mr Robinson is courageous', or 'Mr Atkinson is cowardly' are, if true, statements of fact; but we cannot place one as better than the other without applying some standard; and this standard, according to Hume, cannot in itself be a statement of fact because statements of fact cannot be statements of value.

But there is another case where the word *ought* occurs which I don't think even Hume would dispute its appropriateness or worth. The word *ought* may occur in a different type of sentence in conjunction with the word *if*; for example, '*If* Tim wants to remain dry while walking outside in the rain, then he *ought* to bring and use an umbrella.' Now the reader may object that there are other solutions to the problem; if so he may add additional clauses to the statement, '*If* Tim wants to remain dry while walking outside in the rain, then he *ought* to bring and use an umbrella or wear a hood or ....' This quibble does not adversely affect my case; but I shall just consider the single *ought* for simplicity. In this context, the use of the word *ought* is unequivocally correct. So let us consider what we are doing in this example, and that may help guide us to how we should consider *ought* in its ethical sense.

Firstly, we have stated a purpose: to remain dry; indeed every use of *ought* in this sense is related to purpose. This, I think is quite general: the word *ought* is always linked to the fulfilment of a purpose, and this will include moral uses of the word. Secondly, we have named a subject, *Tim*. In doing so, we implicitly define what sort of being that subject is; we have specified his form. Thirdly, we have assumed that the subject has the power and ability to carry out the action implied by *ought*. '*If* Tim wants to remain dry while walking outside in the rain, then he *ought* to walk on top of the clouds' would work as a concept, but as advice it is useless as Tim does not have the ability to walk on the clouds. Fourthly, the subject need not follow the purpose; we do not say 'If, after slipping off a cliff-face, Tim wants to be subject to the law of gravity, Tim ought to continue falling' again is of little practical use since Tim has no choice in the matter. Fifthly, Tim has a will; he can choose to bring an umbrella, or he can choose not to. The question then is whether Tim's will is in alignment with the purpose. Sixthly, the purpose needs to be in accordance with the action, so we cannot say '*If* Tim wants to remain dry while walking outside in the rain, then he *ought* to eat a deep fried Mars Bar for lunch.' While deep fried Mars Bars may have many fine qualities, keeping oneself dry is not among them. Finally, the purpose needs to be achievable given the powers and abilities of the subject. For example, it is difficult to know how to complete the sentence, '*If* Tim wants to jump unaided over the Empire State Building, then he *ought* to ...,' except perhaps to say something like 'Give up' or 'Find another hobby.'

I should also add that this type of statement need not only be applied to actions, but also to being; for example we might say '*If* Tim wants to *be* virtuous, then he *ought* to *be* loving.' While the first example, related to action, relates to duty system of ethics, then this type of structure is related to a virtue ethics. Equally, we can use an *Unless/ought* statement

as well as *If/ought*, for example '*Unless* Tim wants to get wet while walking outside in the rain, then he *ought* to carry and use an umbrella' is a similarly structured sentence that has a universally acclaimed acceptable use of the word *ought*.

The question we need to answer is whether we can find cases of this 'if/ought' statement where we can remove the 'if clause' without affecting the meaning of the sentence, just leaving us with just, 'Tim ought to bring and use an umbrella.' This would, if true (which in this case it isn't in general), be a moral statement: Tim, simply because he is Tim, ought to act in this way or be this way. So are there any 'If/ought' statements where we can remove the 'if' clause without affecting the meaning of the sentence? It is fairly obvious that there are: when the purpose described by the *if* clause is a tautology, or its contradiction is an impossibility given the powers and abilities of the subject, or Tim, by nature of being Tim, cannot avoid the purpose. For example, 'If a writer wants to fulfil his purpose of writing, he ought to write' is pretty much a tautology; a writer is by definition someone who has the job description of writing; if they cease to do that they are no longer a writer. In other words, the position of a writer is defined by the purpose of writing. So the *if* clause reduces to 'If a writer wants to be what he is', which is essentially a tautology since something cannot be what it is not, so we can rephrase this as 'A writer ought to write,' a statement which I hope is wholly uncontroversial. So if we can find an if/ought statement where the if clause is superfluous and can therefore be omitted, we can reach an ought statement without the if; a statement of morality derived from fact, and we thus disprove Hume's assertion. The only challenge remaining is to find an example where this is possible in general.

Objection:
By a similar argument we can say that 'A murderer, if he wants to be what he is, ought to commit murder' implies that 'A murderer ought to commit murder.' However this is false. Therefore the argument above is flawed. •

Response:
There are two types of purposes: general which apply to all men, and specific purposes, which only apply to some people in certain circumstances. The purposes related to certain jobs, occupations, or situations in life would count as a specific purpose, so for example, the purpose of a teacher to teach or a writer to write would be a specific purpose. Equally, the purpose of a murderer to murder, if there is one, would be a specific purpose. However, to judge whether those purposes are good in themselves, we have to compare them against the general purposes of mankind. So, for example, man is defined as a rational animal; the nature of a rational being is to have an intellect, and by definition that which has the purpose of knowing and understanding is an intellect, and since a teacher and a writer contribute to this purpose these occupations are in general good. There is also a natural purpose (which in part defines what it means to be a living organism) of continuing, maintaining, and sustaining yourself, and if we try to construct forms consistent with this purpose, we will describe the various possible forms of living organisms. By comparing these theoretical forms with the creatures observed in practice, we conclude that any living organism, including men, have this purpose. A murderer is by definition someone who interrupts this purpose by ending a life, and therefore, in the absence of any extenuating circumstances (such as a last resort in defence), the purpose of a murderer is immoral. Thus somebody who commits murder may be a good murderer (in the sense of being good at murdering), they are nonetheless a bad human being, and since being a human being is more fundamental than being a murderer, we can say that nobody ought to commit murder. ∘

The word 'ought', in an 'if/ought' statement follows from the purpose as either the only or one of a number of solutions to the the problem raised. Hume's dilemma, then, can be resolved if this purpose follows from what *is*. Now, to do this, we require two things.

Firstly, we need a firm conception of what the subject of the statement *is*; and from that what its essential powers and properties are. This is, of course, a statement describing the object's form. This is the point where what *is* enters the statement. Secondly, the purpose has to follow directly from the substance of the subject; it has a final cause. That is to say, there is a contradiction between what Tim is if He does not do what he ought to do. Now the example quoted above is not of this type, since it is not essential to Tim's nature that he remains dry, gets wet or carries an umbrella; he remains Tim in any case. But there are other possibilities. For example, we may say that 'If he wants to be a rational agent, Gordon ought to think and to will.' He we have the subject, Gordon, who has the form (among other things) of being a rational animal; the purpose, being rational, and the action which ought to be done, namely thinking and willing. Now the statement is clearly correct; for it is the defining property of rational agents that they have the natural tendency to think and to will. The final causes of a rational agent are thinking and willing; and this implies the formal cause of a being with the power to think and to will. Now we usually construct this sentence in the following order: Gordon, purpose, action. We start with Gordon, we consider what his purpose is, and then we decide how he ought to behave to achieve that purpose. But let us consider the sentence structure in the opposite direction. We start with 'ought to think and will.' Next, we construct the purpose, which clearly has to be 'to act as a rational agent,' since thinking and willing is what makes one a rational agent. Then we can consider what subjects are appropriate to the purpose; rational agents, including our friend Gordon. What is applicable to all rational agents is also applicable to Gordon; so we are free to substitute him in. So we have the complete statement. But since the purpose is part of the definition of what Gordon is, it is superfluous, so we can just omit it and say 'Gordon ought to think and will.'

Let us consider a different example. Firstly, I am going to suppose that there such things as natural purposes, which are related to final causes in a manner which I shall describe later. These are connected to the form of a substance; forms are identified by selecting various natural purposes and then seeing which possible forms are consistent with those purposes, since there is only a limited number of forms which enable a particular set of purposes. For example, the purpose 'acquiring knowledge' is only applicable for those forms which have or could possess an intellect, since an intellect is necessary to acquire knowledge. Thus for a being (let's call him Jon) with such a natural purpose, we can write, 'If Jon wants to fulfil his natural purposes, then he ought to be desirous of acquiring knowledge.' 'Desirous' here refers to his natural inclination and character; it doesn't mean that he seeks to spend every moment of every day in a library, only that he has this general inclination. Of course, without such an inclination, he would not be disposed to study, and thus could only fulfil this purpose by accident, which is most unlikely. This if/ought statement is thus correct. So what if Jon doesn't want to fulfil this particular natural purpose? Since purposes are linked to form, he is in effect saying that he would be perfectly content to be a creature that lacked the associated capacity; indeed he would prefer to be such a being since there would then be no chance of him using that faculty unintentionally. In other words, he is saying that he would be prefer to be something that he is not. But this is irrational, since one cannot be something one is not. So we have, 'Unless Jon would be content to lack an intellect, he ought to be desirous of acquiring knowledge'. But Jon naturally has an intellect, so to desire not to have one is wholly irrational. Since the unless clause is irrational, in effect a contradiction, we can omit it (in the same way that we can omit tautologies in an if clause) and state that 'Jon ought to be desirous of acquiring knowledge.' But this is a moral statement, an *ought* on its own with no corresponding *if* or *unless*.

This argument is, of course, based on several assumptions. Firstly, that a statement that to be content to lack an intellect is wholly irrational for a being whose natural purposes contains the intellectual purposes. This is the same sort of thing as saying that to desire to be something you are not is irrational, something I take to be self-evident. Secondly, it relies

on the scholastic notion of form; in other words, that a being can exist in numerous different states, only some of which are aligned with the natural purposes and some of which aren't or are only partially[5]. The general statement of morality is that the agent ought to seek to be in one of those states in full alignment with one's natural purposes. Finally, we assume the existence of natural purposes. These are related to and ultimately derived from final causes. Thus this approach only works in system of scholastic metaphysics, and because of the close links between form and finality, so the natural purposes are in effect included within the definition of what something is. And this, of course, is the main reason why Hume's is/ought dilemma fails in the scholastic system: the full definition of a substance includes its final causes or purposes as well as the formal, material and (for a particular object) efficient cause; the *ought* is contained within the full definition of the *is*. Indeed, the final cause in scholasticism is the mother of all causes, and it is the final causes which provide the basic foundation upon which the formal and material causes are constructed. So while Hume's dilemma, or Moore's naturalistic fallacy, pose a serious issue for the enlightenment philosophers who accept Hume and Moore's presumptions, to the scholastic philosopher they are an irrelevance.

The scholastic system of interconnected final and formal causes means that Hume got the whole problem reversed. We don't go from *is* to *ought*; we go from *ought* to *is*. We start by considering a set of natural purposes; we construct the various forms that are consistent with those natural purposes by first identifying an individual that has the capacity to fulfil all those purposes, and then using the principle of continuity to fill out the whole form. Of course, we will obtain numerous different theoretical forms for each set of purposes. This leads us to the conclusion that 'If this type of being exists in nature, then it ought to be in one of the states that best fulfils its natural purposes.' Finally, we compare those sets of abstract notions against what we observe in nature to identify which particular creatures belong to which of the pre-calculated forms. In short, the inherent purpose becomes the key element of the definition of the creature. Of course, this type of argument leading to what something *ought to be* naturally leads us to a virtue ethics rather than a duty or consequentialist ethics. This is the basic principle behind natural law ethics.

So how have we fared against Moore? The supposed naturalistic fallacy is based around the idea that the word *good* is both meaningful and undefinable. *Good*, however, is expressed as centre of ethics: it provides the standard by which we judge what we ought to be. So, let us call this version of good the E-good; to distinguish it from other conceptions or definitions of goodness. So (for example) when someone asserts that pleasure is good and good is pleasure, they are defining *good* as something else; not as the moral standard, but something which is equivalent to pleasure. Let us call this the R-good. The challenge comes in showing that the R-good can be equated with the E-good. A R-good is definable in terms some properties or powers or states of its subject. Many people have defined R-goods in various (contradictory) ways. What Moore claimed is that there is no possible definition of R-good that can be rigorously shown to be equivalent to the E-good. People are confusing two different senses of the word good, one based on their arbitrary definition, and the other which they then use to construct their ethics; and this is the fallacy.

Now I quite agree with Moore that we should not equate E-goodness with pleasure, or happiness (as the word is understood in modern English). But his claim is far stronger than this; he claims that it is impossible in principle to leap from the R-good to the E-good; he proves that it is not achieved for certain examples. Once again, Moore's argument has a simple and obvious flaw: he fails to acknowledge the existence of formal and final causes, and in particular the link between the two of them. For, as we have seen, if we express the

[5]Note that Jon can move from one of his potentia to another and still remain what he is; the contradiction at the heart of my argument above is that if Jon desires to be in a state which is not described by any of his potentia - such as one of the possible states of a particular dung beetle or some other non-rational animal real or imagined.

R-good in terms of the final causes of a particular form, then, on pain of contradiction, this must be equivalent the E-good. For goodness (in general) is defined as being fit for purpose, and obviously if one denies final causes and thus intrinsic purposes then this is indeed a mystery. But, as shown in earlier chapters, we should not deny final causes. There is still a jump from final causality to purpose, but, as I shall show later, that jump can be made; it is, after all, much smaller than the jump from structure to purpose.

14.3.6. Objective Ethics without God?

It is often stated that it is impossible to have a consistent ethical theory built from first principles without either directly or indirectly requiring the existence of God. Note that this statement is very different from saying that atheists can have no ethical theory or that atheists are necessarily immoral (and still less that theists are necessarily moral). Nor does it say that every form of theism will be consistent with a reasonable ethical theory. It is quite possible (if this view is correct) for atheists to sincerely believe in a solid, well-founded ethics and in the absence of God. But what is claimed is that the belief in ethics is inconsistent with the belief in atheism; either the atheist's ethical system is ultimately incoherent, or it conflicts with his lack of belief in God. Of course, the inconsistency might not be apparent without far greater thought than is usually applied. People can, and almost invariably do at some level, hold several contradictory beliefs without realising that they are contradictory.

While Hume himself was in general a rather sloppy thinker, many more careful philosophers have accepted that his reasoning was sound in this case. If you agree with his premises then you must agree with his conclusions. So when Hume states that there is no objective ethics, which is a corollary of his is/ought distinction, then to challenge this notion you have to point out one of his premises that you believe to be wrong. Most people, even those who claim to accept moral relativism, in practice believe that at some level ethics is objective, that is, that there are moral rules which a person cannot override by themselves because they happen to have a very different but (to the relativist) equally valid world-view. Suppose, for example, there is a man who claims that rape is morally acceptable, and has a well-developed theory to justify his belief. I think everyone, even the most ardent moral relativist, would react to that man with horror. But why? Because we are not allowed to act against someone else if they believe that it is wrong? But that is an objective moral rule, so if we accept Hume's premises we must deny it. Clearly the rapist will say that we are permitted to do whatever we like as long as at least one of the parties to the action thinks it respectable (he would, of course, to be consistent, not have a thing to say against someone who rapes him – he might resist at the time, but after the deed is complete fully defend the right of the abuser to do what he did). What right have we, if moral values are purely subjective, to say that he is wrong; to impose our own objective rule on him? Because rape is against the law? Several African and Middle eastern countries treat homosexuality as illegal; does that mean that in those countries homosexual practice is immoral but moral in the countries where it is permitted? What right have we to object to the law (either in the Middle East or the West) if there are no objective standards against which to judge the law? And if this law may be questioned, then why not also another law, even one that is presently universally held such as the law against rape? Because the majority of people think that rape is immoral? But there was a time when most people thought that slavery was perfectly justified. Should we then condemn the abolitionists for their gross immorality, in going against the perceived view and trying to bring an end to this practice which was, until they swayed public opinion, perfectly good and decent? And in any case, if ethics is to be determined by the beliefs of the majority – that, as in all of these prescriptions, is another objective rule, which can be reasonably disputed, and therefore cannot be accepted by the moral relativist who practices what he preaches.

So what of some of the general ethical principles proposed by secular philosophy?

- We are to maximise well-being. How do we define well-being in a way that is both subjective and universally applicable? Do we maximise the well being of each individual as they themselves perceive it? But then how then would we settle disputes, where one man's well-being is fundamentally dependant on his playing of music late into the night, and the neighbour's dependant on getting a good night's sleep? How do we judge which claim is better, without proposing an external rule upon them? How do we even measure the relative well-being of each person so we can accurately them, especially if we are not allowed to bring any non-subjective principle into play? And how shall we counter the man who states that we should each maximise his own well being, if ethical theory is subjective?
- We are to avoid causing suffering. But how are we to define suffering in a universal but subjective way, and why should we avoid suffering? If there is a man who enjoys torturing babies, why should we make him suffer by depriving him of this enjoyment, or count the baby's suffering (as after all, the baby is not yet fully aware and cannot truly be counted as a person[6]) as greater than his? If you cannot get is from ought, then the baby-torturer's view of what we ought to do is just as valid as ours.
- That we should do to others as we would have others do unto us? A good rule, in the proper context – but why should we believe that what we want for ourselves is good even for ourselves, let alone for someone else?
- That ethics ultimately emerges from evolution? That the rules we govern ourselves by were tuned by natural selection to optimise the chances of society's prospering? Firstly, this is of little comfort to the modern ethicist, with his determination to overthrow the rules optimised by millennia of societal development, and replace them with something different which by definition must therefore be sub-optimal; tuned, as it were, by the whims of man to hasten society's destruction[7]. Secondly, this fails to answer the question why ought we to benefit society – we are obtaining an ought (the moral rule) from an is (Natural Selection), and so are implicitly saying that Hume got one of his premises wrong. But then, we need to identify which premise that is.
- Nor is it good enough to say that we should each take the principle that seems best to each of us. That is fine, until the different principles conflict. For morality is all about imposing our views on others. When the raped condemns the rapist as immoral, what are they doing but insisting that the rapist conform to the ethical principles of the raped? Governments legislate morality all the time – for example the recent attempts to outlaw speech deemed offensive to homosexuals – what is that but imposing their own form of ethics on another? Or, similarly, a law to outlaw homosexual practice, or slavery, or theft, or anything else. There are many people who espouse this principle – and they are usually the first to criticise and try to outlaw the practice of those who disagree with them.

[6]This is most definitely not my view - being a person is a matter of the form, and the fetus from the moment of conception is of the form of a human being, a rational animal, and thus a person - but it has been proposed by some people (Singer, 1979) who have proposed that a justification for abortion should also be extended to permitting infanticide.

[7]And this is certainly the effect of the past sixty years of social experimentation in Western culture, undoing in two or three generations a thousand years of (mostly) societal progress – the lower than required birth rates, increasing number of children disadvantaged by broken families leading to a growing underclass of youth employability and despair, the increasing gap between rich and poor, the opposition of the parasitical governing and wealth owning classes to the actual producers of wealth, and over-extension of government and unsustainable national debt – throughout history these symptoms have invariably preceded a civilisation's fall by weakening it so it cannot withstand the inevitable economic or military crisis that a stronger society would shrug off (Unwin, 1934).

If the universe is meaningless and without purpose, as many modern atheists state, then the only consistent ethics is nihilism – everyone should do what seems right for them, whatever the consequences or harm to others. Indeed, if Hume is right, then 'harm' itself is ultimately a meaningless concept. As I know of no moral nihilists in practice, I assume that everyone is either illogical, hypocritical, or believes that Hume's argument is in some way wrong.

So which of Hume's premises do we challenge? Of course, there are many different types of atheism, and each of them are built on different metaphysical assumptions. Many - for example the 'scientific' materialism that dominates much of contemporary thought – are explicitly built on the same starting point as Hume's theory. The premise challenged here – that there are no formal or final causes inherent in material objects (i.e. as real features of the substances rather than just useful labels used in expression) – is inconsistent with atheism. The avoidance of Hume's dilemma through natural law ethics invokes God in two ways. Firstly because, as was shown earlier, the existence of formal and final causality necessarily leads to a God who possess both intellect and will, which atheism by definition denies. Secondly because we have required a connection between final causality and natural purpose. If we can invoke God, then this connection is straight-forward: God created everything positive in the universe for a reason[8], and that reason implies that things have a natural purpose. At the level of fundamental particles or inanimate objects, these are the final causes; at the level of more complex living organisms these are what I have been calling natural purposes. The only question then is what that natural purpose is, which we can do by asking what forms are consistent with a consistent set of natural purposes.

The alternatives are divine command and an Kantian ethics derived solely from the axioms of logic. Divine command the atheist obviously has to reject. The Kantian system comes with numerous problems and difficulties; and I will discuss the objections to it in detail below, but here I will just say that it is widely accepted that Kant failed.

In short, there is no ethical theory that is not in some sense built on the arbitrary whim of men that is consistent with atheism. If we suppose there are moral principles, then we have to accept God. The theory of moral principles and atheism are inconsistent with each other.

We thus come to the syllogism

(1) Objective morality is incompatible with all forms of atheism.
(2) A universal morality has to be objective.
(3) At least some universal moral rules exist.
(4) Therefore every form of atheism is false.

By 'universal' I do not mean that precisely the same rule applies in every situation, but that there is a rule that applies to everyone. This rule may be of the form 'X applies in circumstance A, Y applies in circumstance B' and so on, as long as there is no overlap between the subsets A and B if X and Y are contradictory. I also note that by objective I do not mean that there is universal agreement, but that it is in principle impossible to disagree while one's beliefs remain self-consistent and consistent with the natural world (including the unobserved or not-yet-observed aspects of nature).

Of the three premises, I have already argued in depth for the first two. There may, of course, be some form of atheism of which I am unaware which is self consistent, admits final causality but evades the arguments of the previous chapter. But unless this is specified, I think that we have to treat the first of these premises as secure. The third premise is perhaps least secure, though in practice none of us are comfortable denying it. If these premises hold, then the conclusion follows.

[8]By *everything positive*, I mean to exclude those things which are distinguished by a defect, absence or a lack. Since evil is the privation or lack of good, we cannot thus say that God has a purpose for evil; rather that God has a purpose for everything in as much as it is good.

14.4. Virtue, Duty, and Consequentialist Ethics

14.4.1. Virtue based ethics

DEFINITION 14.4.1. **Character:**
Character (in the sense used in ethics) is the natural predispositions of a rational agent. What a rational agent has the tendency to do in a given situation is determined by his character. △

Discussion:
That people have characters is evident from observation. We know that if we put people in the same situation, they will act in ways which are consistent with their own previous actions, but which are inconsistent with how other people behave in that circumstance. □

DEFINITION 14.4.2. **Virtue:**
A Virtue is an individual element of a good character. △

DEFINITION 14.4.3. **Vice:**
A vice is the lack of some virtue in a character. △

14.4.1.1. Plato

Plato's philosophy (or that of his mentor Socrates) is very much the basis of classical Western thought (as well as classical Eastern European thought, to an even greater extent). Of course, it was not originally so. Plato was just one of many Greek and Roman philosophers, and not all were of his school or agreed with him. However, Plato, and his successors grew to be regarded as the best of the philosophers; the clearest in their definition and methods, and they had particular appeal for Christian and Jewish theologians. For though Plato was himself a pagan polytheist in his worship, his philosophy, as developed by the school that followed him, took on a distinct monotheistic tone. Much of what Plato wrote was in accordance with Christian theology in the Bible; so much so that some Christian writers speculated that he had been inspired by Moses[9]. As Christianity rose to prominence in the Roman Empire, so did Plato; the victory of the Church was also the victory of Plato, and since then until the Enlightenment, Plato, and those following him based on similar principles (such as Aristotle) dominated European thought.

The Church was quick to adopt Plato's language as it tried to explain itself to the Greco-Roman world. Of course, that does not mean that Christianity primarily originated from Greek thought. The New Testament, even though Paul was familiar with Greek thought and some of John's writings reflect the Jewish Platonist Philo, is thoroughly Jewish. The earliest major Christian philosopher I am aware of was the early second century theologian, Justin Martyr, whose most important writings (unfortunately only a few survive to our day) date from about thirty or forty years after the New Testament was completed. Justin most likely belongs to the third generation of Christians, one removed from the Apostles. It was not, though, until the opening of the theological schools at the start of the third century, first in Alexandria and then Antioch and Rome, that Platonism started to dominate the Church, and the first truly great Christian Philosophers emerged: names such as Clement of

[9]See, for example, Augustine *City of God*, Book VIII, Chapter 11.

Alexandria, Origen, and culminating in the greatest of them all, Augustine of Hippo. Neo-Platonism was the language used as the early church completed the work of systematising Christian doctrine; both for the authentic Christians and the heretics.

Of course, questions were raised and continue to be raised about this process. Was Plato adapted to serve the needs of the New Testament, or the New Testament adapted to serve the needs of Plato? Many argued that the philosophers went too far; for example in the ancient Church with the case of the condemnation of Origen for his belief in the infinite pre-existence of souls[10], taken directly from Plato[11]. In later times, the extent to which secular philosophy should contribute to theology was one of the questions that dominated the reformation, and still remains to this day, particularly in the divisions between the radical reformed and conciliar churches.

Plato himself, of course, pre-dated these later developments. It is likely that he, as a polytheist, would not have approved of the use that the Christians later put to his work.

He is best known for his Socratic dialogues. These would generally be set as a conversation between Socrates and several other notable members of Athenian society, in various settings, for example at a dinner party or a chance meeting under a tree. How much of the dialogue was genuine, how much represented the views of Socrates, and how much Plato's own work is difficult to determine. The dialogues take on a standard pattern; the conversation begins, turns to a particular topic, each of Socrates sparring partners would present their own views (which were generally well expressed and seemingly convincing), then Socrates, through force of logic, expressed in part through probing questions, would demolish his opponents and establish his own position.

This format has certain advantages; firstly, it allows a certain anonymity for Plato, who never directly need express the position he takes – useful in the Politically charged Athenian culture of the time. Secondly, it allows Plato to effectively express the objections and different arguments that can be made; in fact, it forces him to do so. Thirdly, it allows there to be someone present during Socrates' explanation, who we can identify with, who will raise the sorts of objections we would raise, or ask for further explanation when we might ask that. Fourthly, it allows for a certain degree of artistic expression. Plato's dialogues are just fun to read, unlike (for example) the more systematic but drier Aristotle. Plato's writings display a great deal of literary skill, both in his mastery of numerous opposing viewpoints, the clarity of his thought, his characterisation, and his use of vivid imagery. Of the philosophers since his day, a few have matched or exceeded Plato in the originality and quality of their thoughts; none (perhaps baring Cicero) have approached his skill as a writer. But the dialogue also has a notable disadvantage: the material is not always arranged logically or systematically. Some of the basic premises behind Plato's thought have to be eked out, a little from this dialogue, a little from that one.

So it is with Plato's ethical thought, his theory of forms. The basic idea can be expressed simply: for every type of thing in existence, there exists an ideal form made by God. This is a real object, but exists in an abstract realm; indeed this abstract realm is the only true reality. The objects we see around us in our world of appearances are imperfect images of this form. Existence involves participating in one of the forms; but all the objects around us only do so imperfectly.

Plato's idea of forms is an early answer to the problem of universals; how different objects can share the same nature. His answer is that they are all shadows of the same being[12]. The nature of the shadow, of course, depends on both the object that casts the shadow and the source of light and how the light is scattered off it and the irregularities

[10]Origen, *De Principiis*, 1.7.4; c.f. Acts and Canons of The Fifth Ecumenical Council (Constantinople, 553).

[11]see Plato, *Phaedo.*, which is a discussion of the immortality of the soul and contains arguments for the soul's existence before birth as well as after death.

[12]I here use an illusion taken from the analogy of the cave in (Plato, 380 BC).

of the cave wall onto which it is projected, and in this way there can be differences in the shadows cast by the same object. Thus this idea explains how there are both similarities and differences in the world of appearances: similarities because the shadows are all of the same source, and differences because they are cast by a different light onto a different location in the surface of the cave. Plato obviously also spends some time extolling the virtues of the philosopher, who can emerge from the cave (though in real life through clear thinking rather than literally), and see the true nature of things. It is also important to note that the existence of the objects in the world of appearances depends on the existence of the ideal form. If the archetype does not exist, then none of the shadows can either; but the archetype can exist without casting any shadows. Moving away from the metaphor, we say that the objects we see around us only exist by participation in the ideal form; this is just a more technical way of saying that their existence depends on the metaphysically prior existence of the form.

There is one crucial point to make, which is implicitly assumed in Plato's model, but needs explicit statement for the arguments that follows. Only positive properties have universal forms. So, for example, there are individual dogs because they participate, to a greater or lesser degree, in the perfect form of dogs. If, however, we consider the property of not resembling a dog this comes from a lack of participation in the ideal form of the dog, rather than there being the perfect non-dog existent somewhere in the world of platonic forms. In fact, this is obvious; for there is (if we accept Plato's model) only one way of being a dog, but an infinite number of ways of not being a dog. There therefore cannot be a single ideal non-dog. This is related to Augustine's answer to the problem of whether God created evil: one can only create that which exists, and evil is, in its essence, non-existence.

I have so far used the theory of forms to discuss kinds of material objects, but it is also applicable in another sense to properties. For example, some objects are red, or triangular. We may then suppose that there is an ideal form of redness or triangularity, and red objects exhibit redness by their exemplifying this ideal form; I use the word *exemplify* rather than *participate* to distinguish between forms of objects and the forms of the properties. Objects participate (directly) in only one ideal form, but may exemplify a number of properties.

Now, it is worth reminding ourselves that Plato's theory of forms is an attempt to explain and understand the concept of universals: to describe how distinct individuals can share features in common with other objects of the same species. We therefore should ask if everything gains its existence through participation in an ideal form. This need not be assumed, and indeed it cannot be the case. For example, if a species has only one member, then the species and the individual are identical; indeed this solitary member must be the ideal archetype of the object; since if it were not the ideal form, then there would also be an ideal form apart from it and therefore more than one member of the species. Equally, we can think of relational expressions, such as 'bigger than', 'the same as' or 'participate in', which also cannot be described by ideal forms; since forms give rise to individuals, while a relation necessarily links two distinct beings. Equally, the concepts of 'number' and the like are not associated with forms. We thus have several different concepts in Plato's world; the material objects themselves, the ideal forms of those objects, the ideal forms of the properties, and abstract concepts such as relation or number.

Now, it may be asked how Plato's theory deals with similarities between different forms. For example, we can say that all wolves are imperfect reflections of the ideal form of the wolf; and this accounts for the similarities and differences between individual wolves. But what about the similarity between wolves and dogs or foxes? Now Plato's theory can obviously be extended to account for this: we suppose that there is an ideal form of the wolf-dog-fox, of which the forms of the species are simply imperfect shadows and which gain existence by their participation in it. This form of the wolf-dog-fox would be even more real (or perhaps we should say closer to being truly real) than the ideal form of the wolf; and the ideal form

of the wolf only gains its existence in view of its participation in the wolf-dog-fox; by sharing some of the nature of the real thing.

We might criticise this aspect of Plato's model: does it not lead to an infinite regress? We would have to keep introducing new universals to explain similarities between the previous level of forms. To obtain the ultimate explanation of dogness would thus be impossible; we can say that dogs can be explained in terms of their participation in the ideal form of the dog, but this leaves open the explanation of the dog's ideal form. We can explain this in terms of its participation in the ideal dog-wolf-fox, but this also requires explanation; and so on, ad infinitum. Thus there is no ultimate explanation; and thus no explanation at all. This objection is clearly fallacious, because there is an alternative to the infinite regression; that is that the series terminates with a single ideal form in which all other existent objects participate, whether directly or indirectly through another form. Now it is obvious what this single form is: the perfect existent being, since the one feature that all ideal forms share is that they exist. We need not consider non-existent forms (which is different from non-existent material objects, for a form may exist even if nothing in our 'world of appearances' participates in it), since non-existence arises from a lack of participation in this ultimate ideal form. Now, it is clear that this perfect existent being would be (were Plato's model correct) God, as defined in the previous chapter. Thus Plato's model can avoid the infinite regression, but only if we suppose the existence of God. Additionally, if Plato's view of realism is correct, then there is a single God. Equally, by considering the forms of intellectual beings, we find ourselves looking at the single perfect intellectual being, and this must be identified with God, as the abstract concept of the perfect intellect cannot arise from participation with something which does not share (at least by analogy) its chief characteristic. Similar arguments can be made for the other classical attributes of God. Does this mean that God is (for example) the ideal bad-smelling object? No, for smell is a property, while an intellect is an object. The regression we have looked at here refers to objects rather than properties. Only certain attributes (the transcendentals) are required to be perfectly exhibited in God; these may be identified in that they can only be exhibited in one way rather than many ways; there are many types of smell, so no ideal form of smell in general, but the intellect is defined only by a single attribute: the ability to grasp abstract concepts, and thus only a single way to grasp an abstract concept (although the intellect itself may arise in many different ways), though many ways in which we can fail to grasp it. Within God these transcendentals are (to preserve God's simplicity) equivalent to each other; so to be the perfect existent means that one also must be the perfect exemplar of intellect, of power, and so on. Equally, there is no ideal form of a lack of power, as this is a negative, so we cannot use this argument to demonstrate that God is the most perfectly feeble being. The implication of the existence of a single God is, of course, is one reason why Neo-Platonism was appealing many early Christians, and Christianity appealing to many Neo-Platonists: there was a distinct discrepancy between the religion of the Greek and Roman philosophers and the religion of the Greek and Roman priests[13].

Others criticise Plato by saying that if there is a gulf between the world of forms and the world of appearances, then we, as part of the world of appearances have no access to the world of forms. This, however, presupposes that the distinction between form and appearances can only be made through a Cartesian duality where this gulf between the material and immaterial is present. This is not true for every type of realism; for example in Aristotelian metaphysics substances are the union between the material matter and immaterial form; there is therefore an immaterial aspect of ourself, which has access to understand the world of forms. Equally, we may deduce (within a bounded imprecision) the nature of the forms by careful analysis of particulars.

A further objection arises from the distinction between universals. The supposition is that two universals U and U' can only be distinguished if there are different objects

[13]Augustine, *City of God*, Book VI, chapters 5-8.

participating in them. For example, the universal 'man' and 'featherless biped' describe, as far as we know, the same objects. So how can they be distinguished – for it is clear that they are not the same. This objection is easily answered, since 'featherless' and 'biped' are properties, while 'man' refers to a species. The way objects exemplify properties is different from the way they participate in archetypes; indeed the archetype for man exemplifies many properties, not just being featherless and a biped, yet cannot be identified with these properties; instead the properties arise from the form of man. There also is a second way in which two universals may have the same members, but be distinct: when U' is more general than U, but it happens that although there are other ideal forms which derive their existence through participation in U', say U_2, U_3, U_4 and so on; these happen to have nothing in the world of appearances participating in them. U' is distinct from U because it can have members other than U, even if, by chance, it does not actually do so. It may be, for example, that if there were no gorillas, chimpanzees, baboons, orangutans and so on that men would be the only primates. However, this does not mean that 'man' and 'primate' are synonymous, because it would be possible for there to be other primates than just man.

Equally, some have argued that the property "does not exemplify itself" leads to a contradiction. So, for example, there are objects which have the property of not exemplifying themselves; most objects that we see around us, in fact. These, the argument goes, if Plato is right, can only gain this property by exemplifying an ideal object that does not exemplify itself. Yet that object, the argument goes, must gain the property of non-self-exemplification by self-exemplification; or if it does not have the property of non-self-exemplification, then it cannot be the ideal form of this property – a contradiction. There are objects which do seemingly exemplify themselves, such as the property of being incorporeal. Now this objection again fails; firstly, because 'does not exemplify itself' is a negative similar to a non-dog, and thus naively we do not have an ideal form of this property. Secondly, even if we allow for such an ideal form, this would be the termination of the formal series. Yet such a termination must be in a single concept which does not have this property arise through exemplification of a form but by being the property itself. Most red objects are red to the extent through which they exemplify the ideal redness. The ideal form of redness does not exemplify redness: it *is* redness, the sole member of the species of *the ideal form of being red.* The ultimate incorporeal being does not gain its incorporeality through participation or exemplification but by being itself; thus this being also does not exemplify itself but merely is itself.

Plato's theory also provides us with a basis for ethics. The key question of ethics is how to define goodness; and in this model there is a simple answer: goodness is a measure of how closely an object resembles its ideal form. Some shadows are largely distorted, others truer images. The analogy, of course, now starts to break down: in ethics, we are not primarily interested in our physical attributes, but instead our mental attributes, our character. There is an ideal man; whose character is fixed. Our goodness is measured by how closely our characters conform to this ideal. Of course, our characters are not fixed; we have power, through exercise and training, to improve them; indeed, it is our duty to do so. For our existence comes from our participation in the ideal; and the closer we come to the ideal, the more real we are.

This picture is useful as a simple prototype of a virtue based theory of ethics; and most of the key elements are present. A virtue based theory of ethics pictures some ideal features of the type of being, and then measures goodness in terms of the differences and similarities to that ideal. This assumes, of course, a workable theory of universals: that beings may be divided into certain kinds, which share numerous features, including the measure of goodness. Indeed, in any such classification, we are close to establishing a theory of virtue based ethics; for as soon as we divide the beings into kinds, we present some sort of norm which represents the beings, and from this some sort of statement about what they ought

to be. It is not a completely rigorous link, but certainly the most plausible theories of metaphysical realism tend to be linked to some system of virtue based ethics.

Plato's theory of forms is a useful first step in explaining virtue based ethics, and the theories of universals, but it has too many flaws to be taken too seriously in its original form; I will just highlight a few of these flaws. Firstly, of course, we have Plato's division between the real world of forms and the world of appearances around us. This division seems to be the wrong way round: surely we live in the real world, and the world of forms should be more abstract? To say that the forms have reality seems to contradict the notion that we only deduce them indirectly. Secondly, it does not allow for variation in the ideal. For example, when we look at the ideal form of a man, is he from Africa, Europe, India or the Orient? If (for example) he is African, then that means that all Europeans are necessarily deficient. But common sense tells us that the African, European, and Asians all share things in common; there ought to be, at least in part, a common standard that applies to all. Plato's theory does not allow us to distinguish between accidental properties of the ideal form, which for ethics are matters indifferent, and essential properties, which are important for ethics. Thirdly, we must compare it to the ideas of modern physics, and here again it comes short. In quantum field theory, there are a number of fields permeating throughout the universe; so we have one electron field, one positron field, one up quark field, one muon field and so on. The fundamental particles we see are, in essence, fluctuations or bumps in these fields. (To be more accurate, the fields represent probability amplitudes for the particles; which takes us even further away from Plato's picture. Of course, there may be some more fundamental field, a single unified field of which the different particle fields themselves are single excitations, but this has not yet been shown and my argument can survive the required adjustments if this is true. We may also ask whether the most fundamental picture of the universe is a quantum field theory, but something else, and we cannot know this until we know the most fundamental theory. However, if I were a betting man, which admittedly I'm not, my money would firmly be on the fundamental theory being some type of quantum field theory: all the evidence we have at the moment is pointing in that direction.) Here we have a solution to the problem of universals: that electrons come from the same field is the cause of the similarities between individual electrons; that the fluctuations are of different shapes and in different places allows for the differences. We can build up more complex objects by arranging the fundamental particles in various ways. This theory of universals is sufficiently different from Plato's that we cannot accept Plato as being the full answer to the problem of ethics. The theory of forms needs, at best, significant modifications, and, at worst, abandonment; except, perhaps, as a simple illustration before moving on to more sophisticated approaches.

Then there is the issue that even if goodness is measured through comparison with the ideal, why should we strive to become good? Once again, the arguments of Hume and Moore need to be avoided, and the theory of forms, by itself without the link to final causes, is not, I think, enough to do so.

Committed Platonists would, I am sure, have answers to these and the other objections; but since I am not a Platonist, I will leave the discussion at this point; and turn to what I think is the most well thought-out version of realism, that of Aristotle, to which none of the objections above apply. See (Irwin, 1979, 1995) for a fuller discussion of Plato's ethics.

14.4.1.2. Aristotle

Of all Plato's disciples, history has given one more prominence than any others: Aristotle. Aristotle was unquestionably one of the foremost geniuses of human history. If he has suffered in modern thought, that was perhaps more due to the misuse his writings were given in the late middle ages rather than flaws on his part. It is better to think of Aristotle in the same way as we now think of Newton and Galileo: not the full answer, and he got many things wrong, but an important step in the right direction. The problem is that, after

the rediscovery in Western Europe of his writings in the middle ages[14], people treated him as though he were the final word, rather than emulating his own sense of discovery and trying to improve on him as he tried to improve on Plato and Socrates; and to identify and remove his errors rather than embrace them. Aristotle's works, possibly originally lecture notes for his school, were broad, and covered almost every field of knowledge. His works on logic founded the systematic study of the subject, and baring a few errors and omissions still remain one of its best treatments. They ought to be studied by our own schoolchildren before they move on to more modern representations. He produced important works on biology, politics, poetics, rhetoric, and many others, and his works would not be bettered for millennia, if at all. Of his physics, astronomy and metaphysics I have already spoken in detail. His work was based on firstly observation and secondly the theory informed by his metaphysical principles (though they themselves, I think, were first brought out by observing how the world appeared to work rather than plucked from thin air). There were only a few fields of knowledge where he was surpassed by others of his time, perhaps most notably mathematics, geometry and history; otherwise he was one of the few people in history who knew almost everything that was known in his time.

Aristotle was born to an intellectual family in Macedonia in Northern Greece. After outgrowing the limits of the local education system, he travelled to Athens to perfect his thought in the leading institution of his era, Plato's academy. He left Athens around the time of Plato's death, and travelled around what is now Turkey and the Greek islands, researching (mostly empirically) the local flora and fauna, and perfecting his own thought. Returning to Greece, he established an academy in Macedon, where he taught (among others) the young Alexander the Great and other Macedonian nobles. After this, he came back to Athens and established a school, where he taught for twelve years before being driven out for political reasons. It was probably in Athens that his most important surviving works were written.

Aristotle wrote three surviving works on ethics; the summary Magna Moralia, the shorter Eudemian Ethics, and the longer Nicomachean Ethics. His ethical thought was partly based on his predecessors Plato and Socrates, and partly on his own metaphysics. I do not intend to review all of his work here, but only highlight a few of his important ideas.

As with Plato, Aristotle's work was primarily a virtue based system. The opening words of the Nicomachean Ethics are a useful summary of his starting point:

> Every art and every inquiry, and similarly every action and pursuit, is thought to aim at some good; and for this reason the good has rightly been declared to be that at which all things aim[15].

Thus we see immediately the link between ethics and purpose. Goodness is defined as that which all things aim at; what I have called in this work the final cause (definition 4.5.10). Now, obviously, we have a different hierarchy of aims. Some things we aim for as a step towards some other good; others as part of a craft or trade, and others still for their own sake. The question then becomes to identify the chief good; the most fundamental principle of ethics.

> If, then, there is some end of the things we do, which we desire for its own sake (everything else being desired for the sake of this), and if we do not choose everything for the sake of something else (for at that rate the process would go on to infinity, so that our desire would be empty and vain), clearly this must be the good and the chief good. Will not the

[14]With the exception of the *Nicomachean Ethics* and his analytics, they had been lost to the Latin world alongside the rampant destruction that surrounded the collapse of the Western Roman Empire; but the texts were preserved by the Greek Christians, and from there passed to the Mesopotamian – modern Iraq and Iran – Christians and Jews, from there to the Arabs, when the Jewish and Islamic commentators adapted his work to monotheism; and then into Western Europe through the back door via Spain or Sicily, or carried back by the crusaders.

[15]Aristotle, *Nicomachean Ethics*, Book I, 1.

> knowledge of it, then, have a great influence on life? Shall we not, like archers who have a mark to aim at, be more likely to hit upon what is right? If so, we must try, in outline at least, to determine what it is, and of which of the sciences or capacities it is the object[16].

Aristotle continues by suggesting that rather this science is contained within politics; that the general good of the people is superior to the good of the individual. We next introduce an important concept within Aristotle, though often misrepresented by later readers.

> Let us resume our inquiry and state, in view of the fact that all knowledge and every pursuit aims at some good, what it is that we say political science aims at and what is the highest of all goods achievable by action. Verbally there is very general agreement; for both the general run of men and people of superior refinement say that it is happiness, and identify living well and doing well with being happy; but with regard to what happiness is they differ, and the many do not give the same account as the wise. For the former think it is some plain and obvious thing, like pleasure, wealth, or honour; they differ, however, from one another – and often even the same man identifies it with different things, with health when he is ill, with wealth when he is poor; but, conscious of their ignorance, they admire those who proclaim some great ideal that is above their comprehension. Now some thought that apart from these many goods there is another which is self-subsistent and causes the goodness of all these as well[17].

The problem is that *happiness* is a poor translation of the Greek *eudaimonia* (*ευδαιμονια*), though it is a very common one, and an even worse translation of Aristotle's use of the term, and this, though most are aware of it, is a serious difficulty in our modern ability to understand Aristotle. Many modern philosophers used the idea that happiness was the chief moral good, slapped on their own definition of the term, and proceeded in a direction which Aristotle would have abhorred while claiming peripatetic authority. The chief difficulty is that there is no modern English word which describes the underlying concept. To avoid misunderstanding, I will therefore use the transliteration of the Greek to represent this concept. The question Aristotle proceeds to discuss is how to define and identify eudaimonia. He quickly dismisses the idea that it is to be identified with pleasure, honour or wealth, before discussing and dismissing Plato's idea of good; on the grounds that it defines good in terms of relation while it should properly belong to the metaphysically more fundamental concept of substance, and that it does not allow that the good may be different according to situation, different times, or the differences in the individual concerned. What is good for a soldier need not necessarily be good for a farmer; If there is one ideal form which describes all men, then what is good for the soldier would also be good for the farmer; yet this is surely not the case, as the military and agricultural arts require different skills and abilities. The good, therefore, is not some common element answering to only one ideal form.

Aristotle then turns to his own definition:

> Let us again return to the good we are seeking, and ask what it can be. It seems different in different actions and arts; it is different in medicine, in strategy, and in the other arts likewise. What then is the good of each? Surely that for whose sake everything else is done. In medicine this is health, in strategy victory, in architecture a house, in any other sphere something else, and in every action and pursuit the end; for it is for the sake of this that all men do whatever else they do. Therefore, if there is

[16]Aristotle, *Nicomachean Ethics*, Book I, 2.

[17]Aristotle, *Nicomachean Ethics*, Book I, 4.

an end for all that we do, this will be the good achievable by action, and if there are more than one, these will be the goods achievable by action.

So the argument has by a different course reached the same point; but we must try to state this even more clearly. Since there are evidently more than one end, and we choose some of these (e.g. wealth, flutes, and in general instruments) for the sake of something else, clearly not all ends are final ends; but the chief good is evidently something final. Therefore, if there is only one final end, this will be what we are seeking, and if there are more than one, the most final of these will be what we are seeking.

... The final good is thought to be self-sufficient. ... The self-sufficient we now define as that which when isolated makes life desirable and lacking in nothing; and such we think eudaimonia to be; and further we think it most desirable of all things, without being counted as one good thing among others – if it were so counted it would clearly be made more desirable by the addition of even the least of goods; for that which is added becomes an excess of goods, and of goods the greater is always more desirable. Eudaimonia, then, is something final and self-sufficient, and is the end of action.

Presumably, however, to say that eudaimonia is the chief good seems a platitude, and a clearer account of what it is still desired. This might perhaps be given, if we could first ascertain the function of man. For just as for a flute-player, a sculptor, or an artist, and, in general, for all things that have a function or activity, the good and the 'well' is thought to reside in the function, so would it seem to be for man, if he has a function. Have the carpenter, then, and the tanner certain functions or activities, and has man none? Is he born without a function? Or as eye, hand, foot, and in general each of the parts evidently has a function, may one lay it down that man similarly has a function apart from all these? What then can this be? Life seems to be common even to plants, but we are seeking what is peculiar to man. Let us exclude, therefore, the life of nutrition and growth. Next there would be a life of perception, but it also seems to be common even to the horse, the ox, and every animal. There remains, then, an active life of the element that has a rational principle; of this, one part has such a principle in the sense of being obedient to one, the other in the sense of possessing one and exercising thought. And, as 'life of the rational element' also has two meanings, we must state that life in the sense of activity is what we mean; for this seems to be the more proper sense of the term. Now if the function of man is an activity of soul which follows or implies a rational principle, and if we say 'so-and-so' and 'a good so-and-so' have a function which is the same in kind, e.g. a lyre, and a good lyre-player, and so without qualification in all cases, eminence in respect of goodness being added to the name of the function (for the function of a lyre-player is to play the lyre, and that of a good lyre-player is to do so well): if this is the case, and we state the function of man to be a certain kind of life, and this to be an activity or actions of the soul implying a rational principle, and the function of a good man to be the good and noble performance of these, and if any action is well performed when it is performed in accordance with the appropriate excellence: if this is the case, human good turns out to be activity of soul in accordance with virtue, and if there are more than one virtue, in accordance with the best and most complete[18].

[18]Aristotle, *Nicomachean Ethics*, Book I, 7.

Notice the key place that the 'function of man' plays in Aristotle's thought. Goodness is measured according to how well one fulfils a function. Indeed, I take this as the definition of 'goodness':

DEFINITION 14.4.4. **Good**:
To be good means to be fit for one's purpose (defined in terms of the final cause or natural purpose); that one's form or action should be such that the purpose is most likely to be fulfilled. △

DEFINITION 14.4.5. **Evil**:
Evil is defined as any deficiency or the absence of some good, that is a failure to be in a potentia which is naturally capable of fulfilling that particular being's final causes. It is not thus a substance or quality in itself; it signifies a lack. Note the word 'any' in this definition. Modern usage only treats the worst deviations from the good as evil; we may rightly say that Joseph Stalin was evil, but rarely apply the word to (for example) Tony Blair. This, however, diminishes the severity of even the small deviations from the good; it makes us overlook them. Certainly Stalin was considerably more evil than Tony Blair; but Tony Blair is far from perfect, and that lack of perfection led to many mistakes in his Premiership, which caused a great deal of obvious and unobtrusive harm both to British society and to the wider world (I do not, of course, also wish to deny that he made a few correct decisions as well; nor that Stalin made some good decisions – though fewer in number, and more than outweighed by the severity of his flaws). It is apt to call these imperfections evils; just as it is apt to call my own numerous imperfections as evils. Also note that what counts as an evil depends on the nature of the being. It is an evil for a man to be blind – not that blind men are evil on account of being blind, but that the blindness is an evil (and thus it is good for us to try to cure it, prevent it, and ameliorate its effects) – but not evil for a tulip to be blind, since the ability to see is not a consequence of a plant's set of natural purposes, and its deficiency is thus not a lack in the tulip. △

Both Aristotle and Plato measure goodness in terms of distance from some norm. The difference is that for Plato this ideal is external to the being, while for Aristotle it is a subset of the potentia of the being.

So, for example, a teacher's material cause is a man or woman; the formal cause is the knowledge and educational training he or she possess; the efficient cause is the appointment to an educational institution or an informal agreement with the students; the final cause is to instruct the students so that they may either increase in knowledge or learn certain skills. The purpose of a teacher is to impart knowledge of concepts or skills to the students: a good teacher is one who, all other things being equal, sees his students depart knowing more than when they arrived. A teacher who spends the time fishing rather than instructing remains a teacher (at least for a short while), because he still satisfies the formal, material and efficient causes. However he remains a very bad teacher. Equally, a soldier is a man (the material cause) enlisted in an army (the efficient cause) trained to use weapons and defend himself against weapons (the formal cause) whose purpose is to defend his nation from aggressors, or be an instrument in its aggressive policies. A good soldier is one who, all other things being equal, is most likely to defeat the enemy and win the battle. A bad soldier is cannon fodder. The purpose of a footballer is to ensure that his team scores more goals than it concedes: the good footballer, all other things being equal, does so. Now, of course, the key phrase here is 'all other things being equal'. A good footballer may not win the game, if he plays in an otherwise bad team: what it is saying is that bad team would fare even worse with a bad footballer in his place. Equally, there are matters of fortune; no

matter how skilled the football team, there is almost invariably an element of luck in the final result. So while goodness is defined in terms of being fit for purpose, it is not measured in terms of the result. Rather, it is measured in terms of the form (more precisely, which of the potentia is actualised). A good footballer is one who can pass the ball accurately and shoot forcefully and with precision, has good positional awareness, a strong understanding with his team-mates, has the skills to win the ball when tackling, has good control when running with the ball, can intercept and block the opponent's moves, hide his intentions from his opponents, and, of course, is an able, fast and fit athlete. These are the skills (or among the skills) that usually contribute to the victory: the team that has and uses these skills is far more likely to win the game than one which does not. Regardless of the final result, goodness is measured by the characteristics of the team, part of its form. Each of these individual skills may be treated by itself. The best footballer is the one who best displays the most of these skills.

Thus the function, or final cause, implies a particular subset of potentia within the form which are best equipped to fulfil that function. This might not always lead to success (fortune being what it is), but it gives the best chance of success. This leads to Aristotle's principle 'the function of man to be a certain kind of life;' this certain kind of life is determined primarily by the fulfilment of our rational nature. Now, just as with the sportsman, there are individual elements which make up a great footballer, so it is with the man; these are the moral virtues. The highest good for man, then, is to act and most importantly be in accordance with virtue, and the highest virtues most of all.

The man who has achieved eudaimonia is therefore the man who is fit for the functions described by the final cause of Man; and such a man is characterised by the individual elements which contribute to this type of character, the virtues. Note that this avoids the pitfalls of Plato: it allows for accidental features incidental to the function – it does not require a single archetype; while admitting a degree of universality, there is also room to vary according to circumstance; it does not require Plato's metaphysics in particular the invention of an actual but immeasurable ideal form, nor does goodness subsist in relation but in the essence of being.

Of course, there is more to it than this. One must not only possess the virtues, but also act upon them:

> With those who identify eudaimonia with virtue or some one virtue our account is in harmony; for to virtue belongs virtuous activity. But it makes, perhaps, no small difference whether we place the chief good in possession or in use, in state of mind or in activity. For the state of mind may exist without producing any good result, as in a man who is asleep or in some other way quite inactive, but the activity cannot; for one who has the activity will of necessity be acting, and acting well. And as in the Olympic Games it is not the most beautiful and the strongest that are crowned but those who compete (for it is some of these that are victorious), so those who act win, and rightly win, the noble and good things in life[19].

So, if eudaimonia is a state of the soul, then what of the other goods? Some are prerequisites for eudaimonia; others live in harmony with it. It should also be noted that those who possess this virtuous life also will possess many of the other goods that some aim for in its place. The virtuous man will be happy (in the modern sense of the term), he will have a pleasant life; as pleasure is a state of the soul, and the virtuous man has his soul in the best arrangement. What we find pleasure in depends on who we are. The virtuous man's chief pleasures are in acts of virtue; and of these he is never in short supply. So while not all pleasure implies virtue; all virtue leads to pleasure. Equally, the virtuous man will lead an honourable life; he will be content with what he has, for his chief desire is fulfilled in

[19]Aristotle, *Nicomachean Ethics*, Book I, 8.

his being what he is, and he no longer cares for great wealth (beyond what is necessary) or ambition or the flattery of men. The virtues are durable; as the man who achieves the virtuous life will not depart from it to seek some greater happiness or because of some aspect of his soul which is unfulfilled. Ill fortune does not temper the virtuous man; for his noble character is equipped to survive it. Also, of course, only a living man can be in a state of eudaimonia.

The chief theoretical problem for the ethicist, then, is to identify these virtues and determine the relations between them. The principal practical problem is to train himself and his students to follow the virtues.

> Since eudaimonia is an activity of soul in accordance with perfect virtue, we must consider the nature of virtue; for perhaps we shall thus see better the nature of eudaimonia. ... But clearly the virtue we must study is human virtue; for the good we were seeking was human good and the eudaimonia human eudaimonia. By human virtue we mean not that of the body but that of the soul; and eudaimonia also we call an activity of soul. But if this is so, clearly the student of politics must know somehow the facts about soul. ...
>
> Some things are said about it, adequately enough, even in the discussions outside our school, and we must use these; e.g. that one element in the soul is irrational and one has a rational principle....
>
> Of the irrational element one division seems to be widely distributed, and vegetative in its nature, I mean that which causes nutrition and growth; for it is this kind of power of the soul that one must assign to all nurslings and to embryos, and this same power to full-grown creatures; this is more reasonable than to assign some different power to them. ... However; let us leave the nutritive faculty alone, since it has by its nature no share in human excellence.
>
> There seems to be also another irrational element in the soul – one which in a sense, however, shares in a rational principle. ... No doubt, however, we must none the less suppose that in the soul too there is something contrary to the rational principle, resisting and opposing it. In what sense it is distinct from the other elements does not concern us. Now even this seems to have a share in a rational principle, as we said; at any rate in the continent man it obeys the rational principle and presumably in the temperate and brave man it is still more obedient; for in him it speaks, on all matters, with the same voice as the rational principle.
>
> Therefore the irrational element also appears to be two-fold. For the vegetative element in no way shares in a rational principle, but the appetitive and in general the desiring element in a sense shares in it, in so far as it listens to and obeys it; this is the sense in which we speak of 'taking account' of one's father or one's friends, not that in which we speak of 'accounting for' a mathematical property. That the irrational element is in some sense persuaded by a rational principle is indicated also by the giving of advice and by all reproof and exhortation. And if this element also must be said to have a rational principle, that which has a rational principle (as well as that which has not) will be twofold, one subdivision having it in the strict sense and in itself, and the other having a tendency to obey as one does one's father.
>
> Virtue too is distinguished into kinds in accordance with this difference; for we say that some of the virtues are intellectual and others moral,

> philosophic wisdom and understanding and practical wisdom being intellectual, liberality and temperance moral. For in speaking about a man's character we do not say that he is wise or has understanding but that he is good-tempered or temperate; yet we praise the wise man also with respect to his state of mind; and of states of mind we call those which merit praise virtues[20].

Aristotle's point here is that there are rational and irrational (irrational here meaning not contrary to reason, but merely indicating the absence of the reasonable faculty) elements of the soul, excluding that which is required for growth and so on, shared with the animals, but this is not important for ethics. The irrational element tries to deflect us from eudaimonia; and proof that it is present by observing that all of us are to some extent vicious. But, if there is an irrational element which deflects us from eudaimonia, then since the irrational does not exist in itself but only as a deficiency of the rational; there must also be a rational element that draws us towards it. This is separate from our intellect; the capacity for abstract thought; as the end of this element is not an abstract ideal but something physical. Now if the soul can be divided into the intellect and the will; so too must virtue be correspondingly be divided into intellectual and moral. These virtues are taught rather than being natural to us; and they are trained within us by the formation of habit.

How do we identify the virtues?

> First, then, let us consider this, that it is the nature of such things to be destroyed by defect and excess, as we see in the case of strength and of health (for to gain light on things imperceptible we must use the evidence of sensible things); both excessive and defective exercise destroys the strength, and similarly drink or food which is above or below a certain amount destroys the health, while that which is proportionate both produces and increases and preserves it. So too is it, then, in the case of temperance and courage and the other virtues. For the man who flies from and fears everything and does not stand his ground against anything becomes a coward, and the man who fears nothing at all but goes to meet every danger becomes rash; and similarly the man who indulges in every pleasure and abstains from none becomes self-indulgent, while the man who shuns every pleasure, as boors do, becomes in a way insensible; temperance and courage, then, are destroyed by excess and defect, and preserved by the mean[21].

A virtue is necessarily in the middle between two extremes: a peak has two valleys on either side of it. Of course, we might question about whether a virtue could be destroyed in more than one way; for example if there are several elements contributing to the virtue, then there are numerous ways in which he might have an excess or deficiency: should we be looking for the middle of a volume, rather than the centre-point of a line? This is true for the man in general; he may be a cowardly glutton, a reckless ascetic, a reckless glutton or a cowardly ascetic: the virtuous man is in the middle of all of these. But each virtue is defined as a single element; it can only be in excess or deficient.

This is often seen as the key element of Aristotle's ethics: the principle that the virtues can be identified by being between two extremes. It is sometimes labelled as the golden mean; although it is not the true mean: the best point could well be nearer one side than the other; and where it lies on this line depends on the situation and the type of man or woman we are considering. A soldier, for example, ought to be braver than the accountant: it is his job to advance under enemy fire, while the accountant's job is ill served by it. Yet

---

[20]Aristotle, *Nicomachean Ethics*, Book I, 13

[21]Aristotle, *Nicomachean Ethics*, Book II, 2

both should avoid the extremes of cowardice and recklessness: there are times when even the accountant should risk his life, or the soldier pull back from danger.

I will leave Aristotle's account at this point, since I have what I need. It may be questioned whether Aristotle's concept of the golden mean is accurate: could there not be some virtues which are at an extreme? I am not convinced that Aristotle managed to prove this: to say that most things can differ from the optimum by defect or excess does not mean that everything must do so. Equally, the golden mean does not tell us where precisely the virtue is: courage lies between recklessness and cowardice, but how do we judge where it is on that line?

However, the golden mean is just a guide, rather than the central element of Aristotelian ethics. The centre of his ethics is the relation between what is good and the purpose or final cause of man; that eudaimonia lies in being fit for purpose. We have already seen that this type of construction avoids Hume's difficulties; and so it will form the basis of the ethics we will develop here. This method exhibits similarities to Plato's scheme, in that goodness is again measured in relation to an ideal; but since that ideal is not exhibited in an actual object but rather derived from one aspect of the physical object, its final cause, it avoids the difficulties of Plato's method. For example, we can easily account for indifferent differences between men; if the purpose of women can just as easily be fulfilled by blonds and brunettes, then there is no moral advantage in being one or the other. Aristotle's theory allows for the idea to have accidental as well as essential properties. Equally, it does not rely on the particular metaphysical picture of Plato, but can be adapted for any metaphysics which admits the four Aristotelian causes and the relations between them; this includes my own ideas incorporating the principles of modern physics.

14.4.1.3. Some criticisms of virtue ethics

Objection:
Different virtue ethicists have composed different, incompatible, lists of the virtues. This shows that virtue based ethics is not as objective as it might be claimed; for if it were objective, only one possible list of virtues would be agreed by all. •

Discussion:
This objection may well be applied to ethics in general, as well as to each of the different approaches to it, for many people disagree about many things. Disagreement does not prove falsity. It may arise in a difference in premises, or an error in reasoning. There is certainly no uniform school of virtue ethics; many have come from different premises. Equally, conclusions must be expressed with a possible error. For example, in the Aristotelian school, many have induced the final causes from the formal causes (even though the final causes are seen as metaphysically prior, we can only construct them indirectly by first considering form). This process is not always easy, and in practice will admit to a lack of precision; failure to recognize this can also lead to discrepancies between different scholars. However, it is important to note that while there may be some minor differences, virtue theories when applied in practice also lead to conclusions with a great deal in common. The differences mean that a mistake has been made somewhere; but that there is a mistake in the application does not mean there is a mistake in the premises. Historians disagree about many of the details concerning the Battle of Pharsalus, but we should not thereby conclude that they are wrong in what they agree on, and say that Pompey defeated Caesar. □

Objection:
If it is meritorious to act compassionately, there is surely more merit for the vicious man who acts compassionately despite his nature rather than the compassionate man who just blindly follows what his nature dictates. In this case the vicious man who acts virtuously is judged better than the virtuous man; which surely undermines the whole premise of virtue ethics. •

Discussion:
This objection assumes that merit in virtue based ethics is awarded according to one's actions or the outcome of one's actions. This is false: in a virtue-based ethics, merit arises from what we are, and is only earned by what we do indirectly as a display of what we are. It is character, not act, which determines one's merit and moral worth. The virtuous man is judged better on account of his character and motives; a good act committed for the wrong reasons is just as bad as a bad act committed for the wrong reasons. Equally, one's character determines what we are most likely to do in a given situation; this is not fixed in time. If the vicious man continues to act virtuously from a desire to do good, then he becomes a virtuous man. □

14.4.2. Duty based ethics

14.4.2.1. The categorical imperative

The most common secular version of duty ethics around today was originally due to the German philosopher Immanuel Kant (Kant, 1785; Wood, 2008). Kant was greatly worried by the is/ought dilemma from Hume's writing, and its implications. He reasoned that in principle knowledge could either be derived from our senses, or from what was logically necessary (although Kant is well known for separating the real world and our knowledge of it, and claiming that we can never have an accurate knowledge of the real world). Hume seemed to rule out that we can derive ethics from something we can observe. But that leaves the possibility that we might be able to use some principle which it would be a logical contradiction to deny as the basis of our ethical system. This principle is known as the categorical imperative.

Kant was concerned about what a moral action is, and concluded that an ethical action could only arise from a sense of duty. For example, to give money to charity is not good if you do it just to try to look good to your peers. Kant argued that the only good reason to do something was from a sense of duty, a desire to obey the moral law. Kant believed that all people could in principle be moral, and so he concentrated on the motives (which he called maxims) rather than the consequences. Since we cannot fully control the consequences of our actions, and ethics is primarily about our choices, the consequences cannot be important to ethics. Neither do we have full control of our emotions: we don't choose to become angry, for example, so these are also nothing to do with ethics.

Examples of maxims might be principles such as 'Always tell the truth because people are more likely to trust you if you are honest', or 'Always give to the poor if you feel compassion for them,' or 'Always help those in need because it is your moral duty to do so.' Only this last one would Kant regard as a valid motive.

There are, in Kant's system, certain duties which are binding on everyone, and these are known as categorical because they do not depend on any conditions. 'You have a duty to always tell the truth' is categorical. 'You have a duty to tell the truth if it would benefit your hearer to hear it' is hypothetical, because it depends on certain conditions being satisfied: it tells you what you want to do to achieve a certain goal. It is the categorical duties which form the basis of the moral rules. How then do we know what these are? The most important test Kant gave was universality; an imperative had to be able to be extended to every person. For example, 'Always tell the truth' can easily be extended to everyone;

'Always tell a lie' if universalised would lead to nobody believing anything, and civilisation grinding to a halt. You should not make an exception for yourself. This is seen as a version of the familiar command 'Do unto others what you would have others do unto you.'

Kant also phrased his categorical imperative in different ways, including that we should always treat other people as ends in themselves, never as a means to an end. From this set of simple rules, Kant hoped to build up a complete ethical system.

So what are the weaknesses of Kant's ethics? The first one is whether it provides a unique solution. For example, suppose that there are two different ethical systems which are both self-consistent, maybe even flowing from the same categorical imperatives. So which do we choose?

The second problem is that it can provide some stupid rules. 'Always wear blue socks on a Sunday' can be universalised, yet is clearly not a moral rule. 'Always chop off your son's little finger on his tenth birthday' can also be universalised, and is also clearly immoral. We therefore need additional rules telling us which universal duties are moral and which aren't.

But then we come to the third problem: which set of categorical imperatives should we choose? And how do we know we have selected the right one? We cannot use logic to decide, because all of them are logical. We cannot use our intuition, because how do we know that our intuitions are morally sound, without some more fundamental theory to guide us? We cannot use our senses, because then we run into Hume's is/ought dilemma (without admitting to formal/final causality). In short, Kant has not answered the fundamental issue.

Fourthly, is it truly right that moral actions can only arise if motivated by a sense of duty? Why is duty so important? It certainly seems to me that this principle can be disputed.

Fifthly, this approach ignores one of the fundamental problems of ethics: What makes a good man? It neglects that our characters can change; that we can learn to control our emotions.

Sixthly, the idea of rules without exceptions is troubling. 'Always tell the truth' seems like a good principle, until we meet a Boko Haram bomber in the forest asking for directions to the nearest school. And what happens if, in a certain situation, two of Kant's categorical duties contradict each other, so we have to break one of them? Which one should we obey? What principle do we have to guide us?

Some people are also troubled by Kant's neglect of the consequences of the action. What does it matter if we mean well, if in practice we do great harm?

For these reasons and others, even though a few people still support it, I do not think that Kant's ethics can survive. It does, however, have its good points. The idea that ethics should be universal (i.e. everyone in the same situation should be treated in the same way) is important, as is the idea that it should be self consistent, and these can be used to test whether an ethical theory is reasonable. They just aren't enough.

14.4.2.2. Divine Command

One of the most important religious theories of ethics is Divine Command theory. It has been common in Christianity (though far from universally held), and dominates some religions such as Islam. This states that goodness is defined as something consistent with the will of God. Thus if God wills that we should not commit murder, then because of that committing murder is always an evil. There are several variants of divine command theory, some stronger than others. The strongest says that the only reason that we can consider something as good is because it is commanded by God. Weaker forms might say that there might be other reasons why we might regard something as good, although goodness is always consistent with the will of God.

Obviously, when we want to know what God commands, the first place we should look is the Bible, or whichever religious scripture we happen to favour, which contains numerous comments on which actions and attitudes are favoured by God. The ten commandments,

and the commandments to love God and our neighbour are the most prominent, but there are numerous other commandments as well which ought to be followed, although, obviously, some care has to be taken to determine whether it is just a commandment of the old covenant or is also applicable to the new covenant[22].

Divine command theory has two important advantages for the religious: firstly, its certainty, at least when based on revelatory texts. Other ethical theories are based on premises and reason, and in both places there is room for errors entering into their derivations. It is, of course, possible to misread religious scriptures (and many have), but in principle as long as we submit to the them as the highest authority, the texts which provide the basic moral rules are usually clear cut and don't require much in the way of interpretation. Of course, that has not stopped people who disagree with a particular commandment from trying, for example, by saying that a certain command was directed towards a particular historical context or sub-set of the cases where one might naively expect it to apply. One of the useful tools to argue against this sort of deception is to ask 'If that is what the Biblical author meant, why isn't it what he wrote?' Almost invariably these modified and unnatural readings would have been better expressed in a different language. We should also look at Church history and tradition: how did our spiritual ancestors understand this passage? If there is an apparently new reading of some passage of scripture which nobody (or nobody of note) over the two thousand years of Christians studying the Bible thought of until twenty years ago, then the chances are pretty good that the interpretation has been read into scripture rather than out of it, especially if the new interpretation bears a striking relationship to contemporary moral fashion.

The second main reason why divine command theory is useful is its simplicity: a Christian reads a command in the Bible, and does it. Of course, this is easier said than done, and often we neglect commands in the Bible, perhaps because we skim over the relevant passages; perhaps because we do not think about ethics when we read the passage and thus judge it against our own actions and behaviour; perhaps because we simply don't read the Bible enough. Many people focus on certain commandments at the expense of others. For example, in today's church there is a particular focus on sexual immorality – which is important because this area is where the church is being challenged by secular culture. But are those Christians also firm enough in preaching against greed and in favour of support for the poor – which both dominate much of the Bible, particularly in the Old Testament? On the other hand, there are also Christians who focus a lot on combating poverty – and turn a blind eye to sexual immorality. We need to avoid both these extremes, and make sure we take a complete view of scripture, and not just focus on our 'favourite' commands.

One reason for believing in divine command theory is the following argument:

(1) There is an objective and unchanging standard of goodness.
(2) Goodness can only be defined in terms of conformity to some will or purpose.
(3) The only objective and unchanging will is that of God.
(4) Therefore goodness can only be defined in terms of the will of God, i.e. by what God commands.

Clearly each step in this argument requires further explanation. I think that the conclusion is solid if the premises are true, but the three premises have each been challenged.

Divine Command does have its weaknesses.

[22]The idea of covenants (similar to a treaty between man and God) is central to Christian theology, and there are important distinctions between the old covenant established with Moses, and the new covenant with Jesus, which were each established for different purposes in different situations and thus have different terms and conditions. Only the moral commandments from the old covenant are part of the new, though in a different way: not as an external standard, but the consequence of Christian belief is for the character and attitudes of the Christians to be transformed towards goodness, and one effect of that transformation, and a sign of whether the Christian faith is genuine, is the desire and ever-increasing ability to follow the moral law in its entirety.

Some, of course, ask why we should follow some bronze age law code, written down with the prejudices of the time and related to that culture, with the implication that because it is ancient it is therefore inferior to more modern views on ethics. This rejection is however, naive. Partly because the antiquity of the religious moral codes is actually an advantage for them: they (and consequently other moral systems which have the same conclusions) have faced that most potent examiner of them all, time, and passed the test. The Judaeo-Christian moral code (even if it was frequently abused) gave rise to the strongest and most powerful civilisation the world has yet encountered; the similar Confucian/Buddhist moral code has sustained the great oriental civilisations for an even longer time period. Modern theories of morality have been dominant for one or two generations at most and already society is starting to creak. Indeed, we should not think that the traditional moral code's rivals are new: similar moral precepts have been tried before, and failed before. It is not the antiquity of the traditional moral codes that makes them worthy of attention - but that they alone out of all the numerous moralities proposed throughout history have survived and given birth to prosperous societies. Alternative moral visions, on the other hand, lead to a collapsing society, and thus within a few generations disappeared.

But surely our changing culture and technological progress makes all historical comparisons redundant? No. Human nature has not changed throughout recorded history (just read the ancient Sumerian or Greek or Indian or Egyptian or Chinese literature, and see how little people have changed in character, passion or ambition since that time), and ethics is concerned with unchanging human nature more than changing circumstance. Circumstances only fine-tune the details of the moral code; the broad principles are unchanging.

But most importantly, this objection fails because it is based on a premise which the divine law advocate would most strenuously deny: it assumes that the religious law codes are of human origin, shaped by their culture; the divine law advocate assumes that they are of divine origin, shaping the culture. In divine law theory, morality is determined solely by the intentions of God for mankind. If a man's character or actions are in line with God's unchanging purposes, then that character or action are moral. The question, of course, then, becomes how we can figure out the mind of God. The answer, according to the advocates of divine law, is that we can't, but that God can (partially) reveal his mind and intentions to us. That revelation is what is recorded within the various religious scriptures across a wide range of times and cultures, and from those scriptures we can piece together the unchanging will of God and therefore the unchanging moral code. Obviously this presupposes that the revelation behind the scriptures is divine in origin; but the believer will accept as much as the sceptic that if it is not, then they lose much of their power and authority: they become just one voice among many. But the religious believer is convinced, by whatever evidence led them to the conclusion, that the scriptures are divine; and therefore any insights gleaned from them concerning the mind of God is just as relevant today as it was when they were first written. Modern philosophies are only as good as their premises, and if the conclusions of those philosophies are contrary to the direct record left by God, then those premises must be wrong and it is modern presumption, not bronze-age tradition, which is in error.

So neither their antiquity nor that they disagree with modern sensibilities are enough to dismiss the moral value of the religious writings. Are there any other reasons which can be wielded against divine command theory?

The most commonly cited weakness is the Euthyphro argument discussed in section 14.2.1, although that argument has often been refuted (including in the above discussion).

A second weakness is how we can think of the goodness of God? If goodness is defined in terms of conformity to the will of God, then to say that God is good is meaningless: it is just a tautology.

A third weakness often cited is that the commands are just completely arbitrary. God would have been free to say that child-rape is moral, and then it would have been moral. We then seem to be fortunate that God's commands are generally benevolent. Also, what if

God changes his mind? This is, of course, answered if we accept the Biblical principle that God doesn't change.

A fourth weakness of divine command theory is that it is incomplete. There are many situations we go into day by day, particularly in the modern world, which are not mentioned in the Bible. How then are we to judge the will of God? For example, we might consider moral questions surrounding genetic engineering, surrogacy, or watching too much or the wrong sort of television. A similar problem is what do we do when circumstances force us to choose between two commandments. For example, the Bible emphases both Church unity[23] and the casting out of heretics[24]. But where do we draw the line? Who judges which party is a heretic, or merely mistaken on a less severe matter?

A fifth weakness is what if the religious text is contradictory? For example, if it says in one place, 'Do not eat shellfish'[25] and in another 'Feel free to eat whatever you like.'[26] In practice, this is not as serious an objection as opponents to religion like to claim, because the major religions all have their means of either dealing with contradictions or ensuring that any apparent contradictions are only due to misinterpretation, misunderstanding, ignorance or some some unjustified philosophical prejudice read into one or the other of the texts. In Islam, there is the principle of abrogation: that any later text replaces an earlier text which contradicted it. Central to Christianity[27] is the idea of covenants: that God made various treaties with individuals and groups of people at different times, and that the duties on the people contained within each of these treaties differ from one treaty to another. In this way the legal or moral codes differ, not because of a contradiction, but because of different circumstances.

However, to my mind, the most important weakness of divine command theory is that we have to first show that religious scriptures truly come from God; and describe which of the religious scriptures to accept. A Christian would accept the Bible as the word of God[28], and consequently also its moral code either as the only guide to ethics, or as a test against which all other ethical systems must be measured. But this is useless for those who are unconvinced by the evidence for the divine authority of the Bible which to the Christian seems overwhelming. Clearly, one can only accept the Biblical description of goodness if one first accepts the Bible. For those who don't accept the Bible, saying 'But the Lord says, "Do not commit adultery," so don't' will only generate a shake of a head and an incredulous stare. But conversion to Christianity begins with the conviction that we are guilty of sin[29], and that there is nothing we ourselves or as a species can do to overcome this problem. Clearly, then, knowledge of sin must come before acceptance of Christianity; which means that we need to be able to point to some objective moral standard independent of the

[23]For example, John 17:11,20-24. Although it is important to question what type of unity is being commended. Does it mean organisational unity (i.e. the organisation of the Church must not be broken)? Or does it mean unity of purpose, or unity in essential doctrine, or something similar (i.e. the Church should be united in doctrine and purpose)? My own reading of the New Testament (e.g. Ephesians 4:3, Ephesians 4:13, 1 Peter 3:8, 1 Corinthians 1:10) is that there is a great deal of emphasis on the importance of unity in purpose, doctrine, accepting Jesus' Lordship, mutual support and encouragement both within and between congregations but almost nothing commanding organisational unity. That doesn't mean that structural unity is not important, but only that it should arise as a consequence of doctrinal unity and not in spite of doctrinal disunity. Christians who regard the unity of the Church organisation above a unity of doctrine (on all points regarded by somebody, especially our ancestors who are the most impartial judges we have available, as essential) within the Church are going astray.

[24]Galatians 1:9, Titus 3:10, 2 John 9-11.

[25]Leviticus 11:10

[26]1 Corinthians 10:25; Romans 14:14

[27]And I do mean central; the scope of this concept goes far beyond the ethical applications I discuss here into almost every part of Christian theology except the doctrine of God.

[28]Which is different from the Word of God.

[29]Sin means to miss the mark; i.e. the target of goodness. Anything that is not good about our character or actions is sin, including all the flaws in character or absence of virtue we obtain when we are conceived.

Bible. When approaching those of other religions, how can the Christian claim to them that Biblical ethics is better than the ethics of their own holy books? For example, by the standards of the Christian holy book, particularly the New Testament, the founder of Islam was (if the early Islamic biographies of him accurately describe his actions, including a lack of repentance) a deeply flawed and immoral person, who committed numerous atrocities[30], and who is undoubtedly now residing in the deepest pits of hell, alongside all of his most devout followers. The Islamic sources, on the other hand, claim that he was the most righteous of all men, and is exalted above all others in the heavens[31], and an example for everyone to follow. Thus, if this view is correct, to emulate him by murdering, raping and stealing is a good in certain circumstances, and to refrain from doing so in those circumstances an evil. So which account is to be believed? This is not, of course, a difference of opinion about what Muhammad did, or what his character was like, since both Christian and Muslim turn to the same sources to evaluate that, and if their reading is thorough enough come away with pretty much the same conclusion[32]. One, however, looks at their holy book and calls those actions evil, and the other at their holy book and calls those actions good.

Some external standard beyond divine command is needed to tell which one, if either, is correct. This might be direct evidence for the truth of the scriptures, or it might be an alternative ethical system. In practice, all the lines of evidence strengthen each other.

14.4.3. Consequentialism

Probably the most popular ethical system in the secular world today (at least among those who study ethics), is consequentialism. This states that a moral worth of an action should be judged by its consequences. One particularly important form of consequentialism is known as utilitarianism, which was popularised and defended by people such as Jeremy Bentham and John Stuart Mill (Bentham, 1823; Mill, 1863).

Utilitarianism is seen to be a combination of two theories: consequentialism, and some other principle which allows us to judge what a good consequence is. In Mill's theory, which I shall outline in most detail below, this principle is the overall welfare of society, defined in a particular way. According to Mill, the mistake of other moral theories is that all systems

---

[30]As one of many examples, his unprovoked slaughter of the Jews at Khaybar, the murder by torture of Kinana bin al-Rabi, the looting of the Jewish community's property and enslaving of their women, and the forced marriage and rape of the Jewish woman Safiyya immediately after she witnessed him murder her husband and family - and that's just one day in the life of Muhammad Ibn Abdullah (Ibn Ishaq, *The Life of Muhammad* 511,515; Bukhari 5.59.552; Tabiri 39.185).

[31]Bukhari, 4.52.74; Koran 33:21,68:4.

[32]This is, of course, a little simplistic, and there are different schools of thought and legitimate differences of opinion about the reliability of some of the earlier texts. For example, there are *hadith* which state that shortly after coming to Medina, Muhammad ordered the assassination of several of his critics (five examples – and there are others - are Al-Nadr bin al-Harith (Ibn Ishaq, *The Life of Muhammad*, 191); Uqba bin Abu Muayt (Bukhari, vol. 4, no. 2934); Asma bint Marwan (Ibn Ishaq, *The Life of Muhammad* 995); Abu Afak (Ibn Ishaq, *The Life of Muhammad* 995); Kab bin al-Ashraf (Bukhari vol. 5, no. 4037; Muslim vol. 3, no. 4436; Ibn Ishaq 548)). These are a few of the many actions which critics of Islam point to to disparage Muhammad's character and offer part of the demonstration that he could not have been what he claimed about himself. There is, however, a question about whether these hadith are reliable. Generally speaking, critics of Islam and Islamic extremists view them as reliable, only drawing opposite conclusions: extremist Muslims believing that it is a good example of the justice needed against those who insult Islam; the critics believing that it is wholly unjustified murder (for example, Christians, or at least Christians who follow the New Testament, would be equally opposed to killing anyone who abused Christianity or the Christian God), and yet more testimony that Muhammad did not follow God but an idol of his own subconscious invention. I have seen more moderate Muslims answer the dilemma by questioning whether these particular stories are genuine, as not all of the hadith and accounts in the histories are as strong as others. All these earliest surviving Islamic accounts of Muhammad's life were written from about one hundred and fifty to two hundred and fifty years after the events they describe, so obviously questions can legitimately be asked about their reliability; but neither is there any obvious historical reason why they should be rejected, as the early Islamic scholars tried to be thorough and corroborated their sources. If we reject these, what reasons have we for accepting those parts of the hadith which we like?

are ultimately based on the principle of utility, but other moralists were blind to this fact and so neglected its importance, or introduced errors or distractions.

In the classic form of utilitarianism, the goal of all people is seen as to be happy. Happiness is desirable, indeed the chief object of our desire. This does not come from a theoretical ideal, but our observation of ourselves and those around us. For example, people try to earn money, or fall in love, ultimately because they think that by such means they will gain happiness. The best state of affairs is therefore when everyone is happy. Our goal then should be to maximise the total sum of happiness in the world. A morally right action is one that increases the sum of happiness. A morally wrong action is one that decreases it. Note that it is the total sum of all human happiness that is important, not one's own individual happiness. In this theory, motive has nothing to do with the morality of the action (though much with the worth of the agent). What then should be our motive for adopting utilitarianism? Ultimately a feeling of the unity of all mankind; that we are all essentially equal, and that therefore we should each feel compassion for the happiness of others. Other motives, of course (such as fear of divine judgement, or being praised by our fellow men) are also possible.

Mill divided ethical theories into two classes: the inductive, based on knowledge, and the intuitive, based on our feelings. Both try to establish general rules. He saw his own idea as belonging to the inductive class. Intuitive theories of ethics can be dismissed: firstly it is not clear that we have an intuitive sense of right and wrong (our conscience might merely be a reflection of our upbringing), and secondly because different people's intuitions contradict each other.

Mill believed that the greatest happiness principle underlies every reasonable account of ethics; even of those who would most vehemently oppose it. Certainly, nobody would want to say that it is moral to cause somebody misery.

> The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure, and the absence of pain; by unhappiness, pain, and the privation of pleasure. To give a clear view of the moral standard set up by the theory, much more requires to be said; in particular, what things it includes in the ideas of pain and pleasure; and to what extent this is left an open question. But these supplementary explanations do not affect the theory of life on which this theory of morality is grounded -- namely, that pleasure, and freedom from pain, are the only things desirable as ends; and that all desirable things (which are as numerous in the utilitarian as in any other scheme) are desirable either for the pleasure inherent in themselves, or as means to the promotion of pleasure and the prevention of pain[33].

Mill extended pleasure and pain from the common understanding to include intellectual as well as physical pleasure and pain, and admitted that it might be right to cause pain initially so that a greater pleasure might occur later, for example when we discipline a child so that he can learn to avoid errors which would greatly harm him in adult life. Indeed, the intellectual pleasures he saw as being of more worth than the physical pleasures; thus some forms of happiness (those which those people who have experienced both generally prefer) contribute more to the sum than others. Neither should happiness be thought of as a single simple object. Rather it is composed of small individual parts, with each part being an end in itself, and not as a means to a greater end. But these parts might differ from person to person. For example, virtue is not naturally in itself one of these ends desired for its sake, but might become so if somebody values it. Equally, the love of money, although money is originally a means to be used to acquire other goods, might itself become a source of

[33]John Stuart Mill, *Utilitarianism*, Chapter 2.

happiness; for example if the fact that we possess money without desiring to spend it gives us pleasure. Happiness is not an abstract idea, but a concrete whole; and these are some of its parts. And the utilitarian standard sanctions and approves their being so.

For Mill, there is therefore nothing that is ultimately desired except happiness, since everything that is desired for its own sake becomes a part of what it means to be happy.

Is total happiness unattainable? And if the goal is unattainable, does that not make its pursuit worthless? However, firstly avoiding unhappiness certainly seems to be possible, and this is a corollary to the principle. Indeed, much misery is caused by human action, or is conquerable by human action, and this misery can be avoided, and will be as mankind progresses to build a more civilised and rational world. Secondly, the principle tells us to maximise happiness given the nature of the world, not to reach the theoretically highest level possible. What of those who have rejected happiness, to try to live for a nobler purpose? Such a thing is not, according to Mill, possible: people rejected some forms of happiness, because they believed that by doing so they would receive a more important (though perhaps non-physical) type of pleasure, or because they would spare many others from the misery they are forced to endure.

Utilitarianism is not inconsistent with valuing the worth of a character, it just sees that as a separate question. The goodness or badness of actions is seen as separate to the question of the goodness or badness of the person who performed them. A good man might on occasion do a bad deed; a bad man might do a good deed.

Mill's argument in favour of utilitarianism can be briefly summarised as follows: 1) The only way to prove that something is desirable is to see what people do desire; 2) We observe that people only desire happiness as an end in itself; 3) Therefore the sum total of happiness is most to be desired; 4) What is most to be desired is the highest good; 5) Maximising total happiness is the highest good.

That, in brief summary, is Mill's version of utilitarianism, regarded as one of the best developed accounts (see (Glover, 1990) for a summary of how the idea has developed since that time). So what do we make of it?

The first and most obvious criticism is that the definition of happiness is unclear. Early in his treatise Mill regards it as as a matter of having pleasure (though not just physical pleasure) and avoiding pain. Later on, he seems to define it as having one's desires fulfilled. But in either case, there is an obvious question: why should we regard happiness as the primary good? In the first case, why should we value pleasure above all? In the second case, why should we value the fulfilment of desire above all? In other words, what makes us believe that the desires are good, or choose which desires are good. Indeed, this comes into the problem of voluntarism: if happiness is defined as the fulfilment of our desire, and the maximisation of happiness is the goal of ethics, then how can ethics be the thing that allows us to regulate our desires, and determine which of our desires ought to be strengthened and which ought to be abandoned? Obviously Mill does better than a naive voluntarism, because the goal is the maximisation of all happiness and not just the individual, so moral living does not mean the fulfilment of our own desires, but the desires of all those in our society; but this does not evade the overall problem but simply moves it up a level. In other words, Mill tries to justify his principle to choose the maximal good as empirically obvious, but his account is not really persuasive once we stop to think about it.

Part of the problem is confusion of the meaning of desirable. This word can have three meanings: what it is possible to desire; what is in practice desired; and what we should be desiring. Mill confuses the second of these three senses with the last one. He establishes what people do desire, but it is a leap from that to saying that we are right to desire it. Mill is also guilty of the fallacy of composition: assuming that what is true for a part is also true for the whole. For example, he states that, in an individual person, each component part of happiness is desired in itself, and therefore happiness as a whole is desired in itself. Problems here include that firstly happiness is thus ill-defined, being the sum of various completely

arbitrary desires, and secondly that these desires might be contradictory, meaning that no total sum of them could be possible. The same issue goes when we try to jump from the happiness of an individual to the happiness of all the individuals in society. While Mill provided a method to weight one individual desire as higher than another, the same thing might be needed when we go from the individual to sum of all people. What if what one man's happiness, given what they value, depends on another's misery? How do we choose between them in Mill's system?

It is possible to reduce Mill's error by saying that something is desired by many people is merely evidence rather than proof that it ought to be desired. However, I don't think that this works. Why is it evidence? Could not most people be mistaken? Indeed, we see throughout history that there have been periods when most people were mistaken about many things. For example, for much of human history most people have regarded slavery as natural and good; more recently most people regard slavery as wrong. They can't both be right. Equally, the Romans and Greeks regarded abortion and infanticide as potentially good (perhaps depending on circumstance); during the Christian era in Europe they were almost universally regarded as always being an abomination; more recently abortion is regarded by many as potentially good, and there are people extending the argument for the acceptance of abortion to promote the acceptance of infanticide. Again, both positions can't be right.

Neither is it clear that even if the majority of people alive or within a society desire something, then that desire is in fact good. The Christian would obviously point to the doctrines of the fall and original sin, and say that if many people naturally desire it, then it is quite likely to be evil. That what people ought to desire can be defined in terms of what people do desire thus needs to be proven, not assumed, and I cannot see how any such proof is possible without undermining the whole foundation of ethics that people do frequently desire evil things or good things in a way that is misdirected or exaggerated and thus distorted into evil.

Secondly, the happiness principle is only one of the two premises behind utilitarianism. The second, consequentialism, is just taken for granted without any discussion in Mill's work. This is also troubling; any solid theory of ethics has to proceed from first principles without making this sort of assumption.

Thirdly, Mill's discussion avoids any mention or even definition of goodness in general.

Another criticism of Mill is that he separates what is a worthy person from what is a worthy action. For example, he stated, 'For certainly no known ethical standard decides an action to be good or bad because it is done by a good or a bad man, still less because done by an amiable, a brave, or a benevolent man, or the contrary.' However, this is close to what the Aristotelians do claim. A good action is one that someone in the state of eudaimonia would do in those circumstances. When Mill talks about good people doing bad things, and bad people doing good things, in neither case does he mean someone who is perfectly good, or perfectly bad, but merely that some people are better than others. However, the measure for the good action is the hypothetical perfectly good person, who is by definition incapable of doing something bad.

This is related to the next criticism: nothing in Mill's approach encourages people to be good. Why should people care about the happiness of others? He believes that the usual reasons apply: fear of hell, desire for heaven, an internal feeling for the unity of humanity. But there is nothing in his principle of happiness as the fulfilment of individual desire that states that these should be promoted, nor people trained in the virtues necessary to act well. After all, a feeling for the unity of humanity is only good (by Mill's measure) if people desire it, or people desire others to possess it, if it is in some way pleasurable, or as a means to one of those ends. If, instead, some people desire to just focus on their own individual pleasure, and this becomes an end in itself for them, then trying to encourage them to act contrary to this desire reduces the sum total of Mill's definition of happiness. However, this is surely wrong.

Furthermore, consequentialism in general can lead to horrific consequences. It is very much a 'means justifies the end' ethics, or one that promotes the 'greater good'. These are, of course, the philosophies of tyrants. For example, the easiest way to maximise the general happiness is to kill all the miserable. Once we remove those in pain (who give a minus score to the sum of happiness), then only those in pleasure or neutral will remain. Nobody will regard this as a justifiable solution. Or again, the thief desires to steal, and stealing gives him pleasure, while the owner desires that his property not be stolen. Who is to say that the pleasure of the thief (and whoever he sells the items onto) is not greater than the pain of the dispossessed? And if so, then stealing would often be justified.

Another criticism that is often made is the complexity of deciding a moral action. Before doing anything, one has to calculate how it will effect everybody: no doubt it will make some people happier, and others more miserable, and one needs to total all these effects beforehand to decide if something is moral. This is, however, just too much work for anyone to do before every decision they make. Equally, we can ask how in practice happiness is measured: an essential point in performing the calculation of what the sum total of happiness is. Do we include the animals? Equally, how far ahead are we to calculate? The immediate consequences, the indirect consequences, or perhaps all the consequences of an action to eternity? There is no obvious cut-off point. For example, most people today would rightly regard Hitler's actions as evil. But what, if among the victims of the Holocaust, there was one individual who otherwise would have had descendants and a thousand years from now one of those would have sparked a nuclear holocaust which would have destroyed the entire human race? But by killing that man, and thus stopping his evil descendant, Hitler stopped the destruction of humanity. So the long term good consequences of the Holocaust would thus outweigh whatever bad would have followed from it. So was the Holocaust good? Of course not; but I cannot see how a utilitarian can make this judgement with certainty given that the consequences I outlined might occur. Utilitarians would take this as wholly speculative, which of course it is - equally possible is that one the possible descendants of a Holocaust victim, had he existed, would have been necessary to prevent the now inevitable nuclear destruction. The good, or bad, done by that man surely outweighs any of the suffering or heroics of the mid-twentieth century. The point is, we can't know what the really long term consequences of a decision are, so even if it seems good now, it might be judged as bad in the future once more consequences have unravelled, or vice versa. This, however, is surely untenable. If the action is regarded as good now, it should surely still be regarded as good ten thousand years into the future[34].

A further weakness of Mill's approach is that he defines happiness, the goal of virtue ethics, as that which people desire, but human desires are changeable. What I desired as a child was not the same as what I desired now; what I desire now will no doubt differ from what I will desire in twenty years time (should I live that long). Similarly, what I found pleasurable as a boy differs from the pleasures I feel now. This means that the goal of utilitarianism is constantly shifting, and if constantly shifting then undefined, and if undefined then useless. How is the utilitarian to judge at which stage (if any) in our shifting desires are they good, and at which stages they are evil?

It is also frequently pointed out that a consequentialist ethics can only be feasible if every good could be measured on the same scale. For example, suppose that we can say that giving to the poor produces x units of happiness, and a certain amount of caring for the elderly provides y units on the same scale. We then have to try arrange our actions so that $x + y$ is maximised. But do we compare between the two types of happiness? How do we value one against the other? How do we even value one of them without being arbitrary? The consequentialist needs to be able to directly compare them to figure out

[34] I should note that there are forms of consequentialism, such as rule utilitarianism, which try to avoid this problem. They supplement a generally consequentialist foundation with various rules or maxims which act as guides in difficult circumstances.

how to maximise his standard of goodness, but how does one do so? Is it even right to think in this way, to making the act of comforting someone in distress equivalent to a matter of merely how much money we put in a box? Some things just cannot be measured on a numerical scale, while others can.

However, I think that the most damning criticism of consequentialist theories (including utilitarianism) is that it can violate the principle of universality, namely that it should apply equally to all people. A corollary of this principle is that if two different people in the same situation with the same motivations were to perform the same action, then the moral judgement of those actions ought to be the same. This is not always the case in a consequentialist theory of ethics.

For example, suppose that we consider that causing a cat suffering is (all other things being equal) morally bad, and relieving a cat's suffering is (all other things being equal) morally good. Only my example, and not the objection, depends on this statement, so if you are ambivalent about the suffering of cats, please feel free to substitute some other example in its place (for example, one can replace the cat with a child). Now suppose that we have a cat which suffers from some chronic condition which is not life threatening, but causes it considerable distress. A mad experimentalist discovers a treatment for this condition, only known to him, and prepares a vial which, if released in the presence of the cat, will permanently and completely cure it. However, he also prepares another vial which contains a toxin which will dramatically increase the cat's pain, but still not kill it. He sets these vials up in a box with two radioactive triggers. If the first particle decays (particle A), it will release the cure and destroy the toxin. If the second particle decays (particle B), then it will release the toxin, and destroy the cure. The experiment is constructed in such a way that if the box is tampered with in even the smallest way (baring putting in or taking out a cat) then the cure is destroyed and the toxin released, and the cat is unable to affect the outcome from inside the box. Now, our mad experimentalist then finds a number of people to act as subjects in his experiment. They are given a choice: either they can put the cat into the box, and seal it, and wait for an allotted time of their choosing before taking it out, or they can decide to leave the cat out of the box. If the cat is kept outside of the box, then it will continue in its condition of moderate suffering. If the cat is placed in the box, then there are three possibilities: either particle A decays first, in which case the cat is cured, or particle B decays first, in which case the cat's suffering increases, or neither particle decays in the allotted time, in which case the status-quo is maintained. Now this experiment is repeated with two different people (who are identical in every relevant detail), X and Y, and two different cats (again, identical in every relevant detail). They each choose to put the cat into the box. X's cat receives the cure, and Y's cat is exposed to the toxin. If the moral worth of an action is judged only by its consequences, then we would have to say that X's action of putting the cat in the box is good, and Y's similar action is evil. But which of the particles decays first, since it is a quantum event, is not just unpredictable in practice, but also even in principle: it is an indeterminate event. It not just happens not to be known, but cannot be known or predicted in advance or in any way controlled. There is no difference whatsoever between X's action and Y's action. There is no difference in their motives. There is no difference in the circumstances surrounding their deeds. Yet we are asked to judge one as good and the other as evil, if a consequential ethical theory is applicable, where the only way to judge the moral worth of an action is its effects. This surely is untenable. My example is only intended as an illustration of a more general point: consequentialism, if it is to deal with people equally, requires a deterministic physics.

This argument does not only apply to quantum systems. Even in a deterministic system, our awareness of the circumstances can only be known to within a certain precision. It may be that X and Y come into some situation, each with the best possible knowledge of the situation but which nonetheless contains some uncertainty. They both make the same calculation, perform the same action, but on account of the minor difference in the real

situation (though their knowledge of the situation was identical), X's action turns out to increase happiness and Y's to decrease it over the long term. In a chaotic system, where minuscule differences in the initial conditions can lead to dramatically different final results even in a deterministic system, such a scenario is certainly possible. So why should we praise X and criticise Y when there is no difference between their knowledge, motivations, actions or measurable difference in the circumstances? Yet that is what consequentialism implores us to do. However, the strongest form of this dilemma arises on account of the indeterminism of quantum physics. In a fully deterministic and knowable world it isn't a problem. Mill himself could not have known about this; in his life time there was no reason to doubt Newton's theory, and like many philosophers he did not realise the significance of experimental errors and in any case chaos was also not understood. In a deterministic theory where there is perfect knowledge, this is not a problem, since the same actions always lead to the same consequences. But now we know better. This argument against consequential theories of ethics is another example of how modern physics can affect philosophical issues in perhaps surprising ways and surprising areas of philosophy which one would naively expect to only be very loosely connected with the structure of physical law.

How are we to solve this dilemma? One approach is to judge the action more on an assessment of what could happen. So, for example, in the case of the cat, one can compute the probability of each particle decaying first, judge the relative worth of the release of the cure and the toxin, and balance all this to decide the relative moral worth of of the action. So if on average, the cat is to get better, then it was better to put it into the box, and if on average it would get worse one should keep it out of the box. But this is not a consequential theory of ethics but a virtue theory. The moral worth of the action no longer depends on the outcome, but on the character and ability of those who act: the awareness of the situation (the virtue of good observation), our ability to compute the odds of each outcome (the virtue of wisdom), judgements of the weight of good and evil (the virtue of prudence), and a desire for the cat's good (the virtue of agape).

Nor, I think, does it suffice to put the blame onto the mad experimenter. For one thing, the actions of X and Y have moral consequences, and thus ought (under consequentialism) to have moral worth; the villainy of the experimenter cannot affect the separate judgement of their own actions. Secondly, we may well (with enough ingenuity) also be able to construct similar examples without any experimenter, but just the blind forces of nature placing X and Y in their dilemma.

Nor is it good enough to say that the universe is deterministic to a very good approximation. Consequential ethical theories demand that it is wholly deterministic; a fundamental ethical principle is either right or wrong; and 'very nearly' is still wrong. This is particularly clear given that there are competing ethical theories which wholly avoid this difficulty.

We thus have two choices, either to abandon consequentialism in these situations, or to abandon the principle that people doing the same thing in the same situation should be judged in the same way, and this second possibility seems to me to be untenable. Note that these examples are no barrier in a virtue theory of ethics, because there one's moral worth is a matter of one's character, and not the consequences of one's deeds.

Now, of course, these examples are somewhat contrived. Can we just say that consequentialism is invalid in these situations, but a valid ethical theory in other cases where our predictions can be reliable? No: a general theory of ethics should be applicable in every situation. We would need some higher theory which tells us in what circumstances utilitarianism applies, and in what circumstances it doesn't. And this higher theory cannot be consequentialist, but something else: either duty-based, or virtue-based, or some other principle. But if that is the case, then we should be teaching that other more general theory as the fundamental ethical theory. Utilitarianism, if it applies in any circumstance, would just be a model which happens to get things right in certain circumstances, and fails in

others. But the more general theory would get things right in more, and perhaps all, circumstances. But if this is true, then one can dispose of utilitarianism altogether, since one could reach the same conclusions from the fundamental theory without reference to it, and not risk using utilitarianism in the circumstances where it doesn't apply.

The best metaphysical approach to quantum physics that preserves determinism is the Everett many-world scenario, and while I don't believe this interpretation to be valid, even so it offers little help to the utilitarian or any form of ethics. For here, with every decision we make, reality branches into different streams. Thus, from the perspective of the multiverse, for each decision every possible consequence occurs in one universe or another. This means that judging moral worth on account of consequences is difficult. Suppose that I have a decision to make now. The consequences of that decision, however, will no doubt depend on further decisions made by either myself, others, or just the natural workings of matter. So there will be future branchings of the universe which affect how the moral judgement is made. If the consequences of my decision turn out to be good, then necessarily in a parallel universe I would have a brother who would make the same decision but the consequences will turn out to be bad. But because the universe is going to branch out in the future, I can't make any judgements now. One of my many-world descendants will have the consequences of the action turn out to be good, and the other parallel version of him will have the consequences of the action turn out to be bad. This means that my action right now is neither good nor bad, or perhaps both of them, but in any case it is impossible to give it a moral status under utilitarianism. This particular point isn't something strictly against utilitarianism: most ethical theories would struggle with the Everett scenario. But, in any case, whether we take the deterministic or indeterminate metaphysical approaches to quantum physics, utilitarianism has a problem.

Mill's version of utilitarianism is not the only consequentialist ethics. For example, there are those who replace his principle of maximising happiness with some other principle, perhaps minimising suffering. Others attempt to combine utilitarianism with duty ethics, giving certain rules which in general will increase happiness (or whatever the principle is which is adopted). However these methods suffer from the same difficulties as Mill's.

I therefore do not think that consequentialism is by itself viable, except possibly as part of a wider system of ethics. Indeed, I do not think that there is any good in the theory that cannot be established from a well developed virtue ethics. Mill believed that all viable ethical systems reduce to utilitarianism; I think the reverse is true, and any viable form of utilitarianism reduces to virtue ethics. Firstly, virtue ethics is required to provide us with the underlying principle: what is the good that we should be maximising? Secondly, virtue ethics is required to ensure that the dispositions of the individuals in society are capable of pursuing the good, by training them in character. Thirdly, if happiness is defined as having those things we desire in ourselves fulfilled, then why stop there? Why not also ask whether we should arrange the desires of the individuals in society so that as many as possible are happy? For example, if one person desires to rape and another to not be raped, then they can't both be satisfied. However, if they both seek to live together in a consensual marriage union, or to stay apart, then both their desires can be simultaneously satisfied. One set of desires can lead to a greater total happiness than others. So we can not only choose to maximise the happiness of the people within a society not only by tuning what actions are performed – and thus the consequences of the actions, but also by tuning what the individual members of the society desire. Indeed, this is far more important, since unless the people's desires are in harmony, it is impossible to make a maximally happy society; but if the desires of people are in harmony, then it is easy to do so. However, once we start judging one set of desires of ends in themselves to be morally better than another set of desires, then we have moved away from consequentialism and into discussing people's character, i.e. a virtue ethics. Thus to maximise happiness, we need to act to encourage the teaching and promotion of virtue ethics.

Now, this idea that good character is one that allows society to live together in harmony is part of natural law ethics (see section 14.5). Mankind is defined in part as a social animal; therefore one of our natural purposes is to live as part of a society, that is people working together for the benefit of them all; and thus harmony is a necessary constituent of a good society (since a community in conflict doesn't work for the good of everyone), and thus the capability to live in harmony with others is part of the natural purposes of an individual man. A good man is one who is fit for purpose, i.e. he has the character to live in perfect harmony with others. Thus the desires of each individual in a good society are such that as many of them as possible can be simultaneously fulfilled. Thus this better version of utilitarianism can be incorporated as part of natural law ethics; but it is only a small part, since living as a social animal is only one of our natural purposes.

14.4.4. Summary

These three basic approaches to ethics need not, of course, contradict each other in terms of their applications. Indeed, if they agree in their basic premises, as far as the different approaches are able to, we might expect them to agree. However, frequently contradictory conclusions arise from applications of the three schools, and in this case we will have to judge between them, and decide which to embrace and which to avoid.

Of the three principle approaches to ethics, I have no doubt that consequentialism considered by itself is by far the worst. The reason is straight-forward: it implicitly relies on a virtue based ethics, but does not in practice consider its foundations. Suppose, for example, that we choose to act to maximise the greatest good: we need some consistent and objective definition of goodness. Before we can ask what is good for a man, we need to consider goodness as it applies to a man; from which we derive what is a good man. But this leads most naturally to a virtue based system. Equally, the principle of least harm requires an objective definition of harm, which can only arise by consideration of man as a species; a task undertaken best by the virtue ethicist. For example, is it always bad to cause someone pain? However we answer this question, we need a reason for doing so; and this reason cannot be provided by solely considering consequences of actions. Secondly, to be successful, a consequentialist society must have its inhabitants schooled in the virtues. For example, they must be able to distinguish good from evil: the virtue of prudence. They must have the courage to do what is good: the virtue of fortitude. They must be able to control their natural inclinations when they would prevent them from acting as they ought: the virtue of temperance. Above all, they must desire goodness: the virtue of agape. Without training its citizens in virtue, the consequentialist society will ultimately collapse into immorality; thus the utilitarian is compelled to teach people to be virtuous; but this requires the understanding and application of virtue ethics, and the original consequentialist beliefs become redundant. In this sense utilitarianism is self-undermining: the one who wishes to maximise the good is forced to adopt and train others in a virtue ethics. However, focusing only on maximising utility (or the good) without providing people with a clear understanding of what good is, the ability to carry out that good, or even any good reason why they should adopt this seemingly arbitrary rule (arbitrary in both the existence of the rule and in the details of its definition) ultimately leads to a society where many people will be primarily motivated by selfishness, minimising the total good. But teach people to be virtuous, without mentioning the maximisation of utility, and you will have both. Thirdly, whatever is present in consequentialist ethics may be deduced from virtue based ethics, if the virtues include agape. For agape is the desire for good, from which it follows that people abundant in this virtue will seek to do good. However, not all of virtue based ethics may be obviously deduced from utilitarianism; for example a virtue based ethic which excludes agape (if correct) cannot be derived from utilitarianism. The utilitarian assumes the goodness of agape before showing it. As in all things, the premises which allow the most general conclusions should be preferred. Fourthly, it is easier to make a mistake in

utilitarianism, since it does not start from first principles but starts from some arbitrary definition of goodness before it considers the nature of man.

This does not mean that consequentialism is wholly flawed; merely that it is flawed if we try to consider it in isolation of or as a replacement of a virtue based system. Utilitarian thinking might be useful to help elucidate some aspects of a virtue based ethics, or to reach some conclusions; but virtue theory is more fundamental.

14.5. Natural Law Ethics

The name 'natural law' might be confusing, since *law* might suggest a duty based ethics, whereas in reality it is virtue based, being basically a refinement and perfection of Aristotle's system developed during the high middle ages. The phrase is used to distinguish it from 'divine law', i.e. those commandments and virtues derived from revelation. Natural law is instead derived wholly from reason combined with scientific observation of humanity. Natural and divine law are not in contradiction, but there are aspects of divine law not derivable from natural law, and some of the conclusions of natural law are not contained within divine law. Given in roots in Aristotle, it should be clear that no mention of the Bible is required in the derivation of natural law, even though it was devout Christians who perfected the ethical system.

Like Aristotle's system, natural law ethics defines goodness as being fit for purpose, and states that, although purposes can be directly specified by an active will, there is another sense of purpose which is not (directly) based on an external will, but arises from the directionality inherent within final causality and entwined in the very essence of matter, and is thus accessible to natural science. The natural purposes are more fundamental than form, and define the nature of the being. Unlike Aristotle, the natural law ethicist does not focus primarily on the intellectual and social aspects of man, but treats his sensory (animal) and vegetative (living) aspects as equally important in framing the moral rules. There is less of an emphasis on the golden mean in identifying virtues, and more of an emphasis on identifying the purposes of man, and the virtues which can be derived from those purposes, and the actions which are consistent with such a virtuous character. Natural law concentrates on those purposes which can be determined through natural philosophy and science, and the virtues which are needed to fulfil those purposes. The Christian would say there are additional purposes of man not accessible to natural reasoning, but which we only come to understand through divine revelation, for example man has the purpose of glorifying God[35]. Thus natural law can only provide a subset of Christian ethics, albeit an important subset. It is not my intention to go beyond natural law to a full discussion of Christian ethics here; but only build my discussion on natural purposes which are accessible through the combination of reason and observation of the physical world. Although inspired by the medieval scholars, I diverge from them in various details, and the presentation here should not be thought of as a description of any particular view of natural law ethics except my own.

CONJECTURE 14.5.1. *The final causes of an object determines the possible formal causes of an object.*

Proof:
The final causes of an object indicate an inherent tendency in an object towards some set of goals or ends. Now, clearly not every possible type of object is consistent with a particular set of final causes. For example, the purpose of a car (its natural tendency if one switches on the ignition, puts it in gear, and applies the accelerator; and also in this case the purpose of its designers) is to travel along roads under its own power (albeit by burning some fuel or through some other source of energy and that most cars require to be directed by a driver) in order to transport a small number of people and a little luggage within it from one place

[35]Isaiah 43:7.

to another[36]. A pram is not consistent with this final cause, since it is unable to move under its own power, without somebody pushing it; a bus is inconsistent with this final cause because it can transport a large number of people. Some cars may lack the capability to move (e.g. due to mechanical failure), but that still does not abate its natural purpose, because purpose and capability are not the same thing; it merely means that the car is damaged in such a way that its final causes are frustrated. The damaged car is still of the same form as the working car, just in a different potentia; and the natural purpose relates to the form as a whole and thus every potentia within it. Thus only a limited number of possible objects can fulfil this final cause; in particular they must have an engine (or some equivalent), doors, wheels, and so on. Thus the final cause restricts the number of possible forms of the object. It does not completely distinguish between different types of being, and we still may add additional distinctions on the basis of form. But the form of an object must normally (baring defect, disorder or damage) be able to fulfil the purpose entailed in the final cause. In this sense, the final cause is metaphysically prior to the formal cause: the form is (at least in part) determined by the final cause. ■

Discussion:
Note that this is not entirely reversible. One cannot directly construct the final causes from the form, because an object may be suitable for more than one purpose. It might accidentally be suitable for purposes that are not part of its final cause. For example, a laser pointer can also be used to throw at people even though this is not its purpose. However, it is nonetheless possible to deduce (though not necessarily with perfect certainty or precision) the final cause from the form; for example, we can see that the final cause of a laser pointer is to be something which we can easily carry in our hand and which emits a correlated beam of light, creating a small dot on a surface, as only laser pointers have this power and all laser pointers have this power; whereas many things which are less complex than laser pointers can be thrown.

A corollary of this is that at least a substantial subset of the arrangements of matter described by a being's form must be consistent with its final cause. Note that not every arrangement of matter which is part of the form need to be consistent with the final cause, as shall be discussed later. □

Comment:
We note that some things are defined not just in terms of their structure or material composition, but also in terms of something else. For example, the difference between a living being and a corpse is not a matter of the way the matter is arranged, but whether they have certain tendencies or purposes. A living being, for example, is often defined at least in part in terms of its tendency to grow, reproduce, and so on. A rational being is often defined at least in part by its tendency to think and grasp abstract concepts.

There is an important difference between capabilities and natural purposes. For example, there are two distinct ways in which we might define a rational creature: either a creature that has the capability of rational thought, or a creature that has the natural tendency towards rational thought, whether or not that purpose is realised in practice in the creature's capabilities. The creature ought to be defined in terms of tendencies rather than its capabilities.

For example, let us define the genera of rational animals, which includes humanity (and may be restricted to humanity), and consider the case of an infant. If we interpret the definition as meaning 'an animal with the capacity for rational thought', then the infant

[36]Some may object to this example, citing that a car is an artefact rather than a substance; its purpose is given it by its human designers rather than the inherent nature of its parts. However, the idea behind the illustration also applies to natural substances; natural and artificial purposes play the same role in their relationships to both form and goodness.

is not a rational animal, because it lacks this capacity (having not yet developed it). If, however, we interpret the definition as meaning 'an animal whose natural purpose includes to think rationally', then the infant can still qualify. It is just that the purpose is blocked because the child is not yet mature. The infant still has the same tendency, but is not yet able to achieve it because it has developed insufficiently. However, we know that the adult is a rational animal under either definition, and also that the adult is of the same form as the infant (as is obvious from the earlier definition of form). If the genera of rational animals is to be defined in terms of capability, then we have a contradiction, since the same form would describe an individual that was both a rational animal and not a rational animal. However, if we define 'rational' as meaning that a the being of that form has the natural purpose of thinking rationally, then there is no contradiction. Defining a being in terms of natural purpose is thus far more reasonable than defining a being in terms of their capabilities.

To take another example, an animal is in part defined by either having the capability or the tendency towards self-movement; movement from one location to another (this is one of the ways in which animals are distinguished from both plants and inanimate objects, so it is certainly part of the definition of an animal). Every human is an animal. Now let us consider a disabled person. This person lacks the capacity for self-movement (in the sense described). If mankind is to be defined in part by having the ability to self-move, then we would have to conclude that the disabled person is not wholly human, because he lacks this ability. If, on the other hand, we define mankind as having the natural purpose of being able to move themselves, then the disabled are not excluded. They still have the purpose, but are not able to achieve it through no fault of their own, but instead because they are prevented from doing so by their injury or illness. Indeed, this gives everyone else the moral obligation to help them to try to overcome their disability, and achieve what they cannot do naturally through artificial means. Capabilities relate to one potentia within the form, so to say that a being has a capability is to say that the actualised potentia has that capability; purposes apply to the form as a whole, so to say that a being has a particular purpose implies that some potentia within the form is able to fulfil that purpose. Since substance is defined in terms of the form rather than an individual potentia, and the definition is related to the substance rather than any particular potentia, the definition of a being should be derived from natural purposes rather than natural capabilities. □

DEFINITION 14.5.2. **Natural purpose**:
A natural purpose is an accumulation of final causes that lead to a macroscopic directionality or tendency in the actions of a complex being. △

Discussion:
We know that every being has a final cause. The final causes exist at the level of physics; in a living organism each of the processes taking place in each cell has its own final cause. The natural purpose is firstly a description of the combined effect of those final causes on the organism as a whole. A natural purpose is thus a group of final causes which can and should be considered together for their combined effects; for example those causes which are related to the processes which allow a living organism to absorb nutrients, or a rational organism to think. Final causes relate to the small microscopic details; natural purposes the large macroscopic effects. The distinction between natural purposes and final causes is just one of convenience when the number of final causes becomes too large to treat each one individually. The same approach is used in physics: for systems with a few particles, we can treat each particle individually and exactly; once the number of particles becomes too large, we need to instead treat the system statistically and restrict our measurements to average values of the ensemble. The natural purpose is to the final cause what the trajectory of

a gas is to the individual final causes of individual gas molecules. These are different in kind: we can discuss the pressure of a gas but not of the individual molecule; the individual molecule's coupling to electromagnetic radiation but the refractive index and absorption spectrum of the gas. So the average properties of the ensemble need not be merely the same as the properties of the individual parts, only on a larger scale; they could be quite different. So it is for the natural purposes as they relate to ethics.

The second difference between the natural purposes and final causes is that natural purposes describe the being as much as possible isolated from its external environment – it refers only to the internal processes of the being as much as possible. 'As much as possible' means (for example) that we should include (for ourselves) eating, drinking, and breathing, and being in air at a suitable temperature range, since they are necessary for us to continue to exist; it excludes, for example, interactions with external agents such as viruses, toxic chemicals, and inappropriately directed machetes, and anything which is not essential for the continuation of the being. For example, lock a man in a room with an inexhaustible supply of sufficient nutrients, water, air, and a comfortable temperature, and nothing else, and he can continue for a long time; exclude any one of these and he will quickly wither. Give a man the nutrients, air, water, a comfortable temperature and a cold virus, and we have given him more than he requires: we can take away the cold virus (though not any of the others), and he will still prosper. We should in this way reduce our consideration to the minimum necessary of external beings for the internal functions to continue to operate. A natural purpose therefore describes the development of the being as it would normally occur in the absence of any external agent preventing the development. A natural purpose is thus a specific set of internal final causes which at the level of the organism are directed towards a common goal. Every being has final causes, but not every being has final causes which can be grouped together into a natural purpose so defined. But it is obvious that some, more complex beings, do. If you want to see one example, then just take a look in a mirror. Natural purposes are more convenient than final causes for the sake of definitions it is easier to describe somebody as rational than to list all the individual processes that take place in the brain which allow that rationality to be manifested. Nor is there necessarily only one set of final causes which lead to a particular natural purpose; there may be many ways in which it is theoretically possibly for thought to be achieved, and our brains are just one example of those ways.

Now, for things to have a natural purpose implies that they must (at least with normal development) either have the capacity to fulfil that purpose or to grow to possess that capacity. In other words, somewhere in their form is a particular state which can at least in part fulfil that purpose. In fact, the relationship between final causality and natural purpose is stronger than this: if certain potentia are actual, then the being will fulfil that purpose unless actively prevented from doing so. If you place nutrients in the digestive system of an animal, then they will be absorbed, unless the stomach or intestines are in some way damaged. This is because animals, like all living things, have the natural purpose of absorbing nutrients, and in a healthy animal the digestive system is the means through which that purpose is fulfilled. So once we identify purposes, we can subsequently (in principle at least, if not necessarily in practice) determine the various individual final causes which allow a being to fulfil that purpose, and from the final causes determine the numerous different forms which are consistent with those purposes. By comparing those theoretical forms against what we see around us, we can identify the purposes of various beings which do exist. So the process of identification is to go from a set of purposes to hypothetical forms (by deduction) to actual forms (by comparison) to individual beings (by specification). In this way, we can link individual beings and purposes.

Natural purposes should be distinguished from intellectual purposes (which is what we commonly think of when we use the word 'purpose'). An intellectual purpose is the goal of an intellectual will, for example, I might desire to buy a house or give money to the poor,

and comparing whether my actions frustrate or further that purpose will determine whether they are good in that sense. A natural purpose, on the other hand, does not directly arise from a will (and the will of God only indirectly), but from the inherent tendencies of a being, as they contribute to the definition of that type of being. So my seeking to buy a house is not a natural purpose, because it is not inherent to my nature as a rational animal that I should buy a house, but rather it arises from a conscious decision; my continuing to endure is a natural purpose because it follows from the definition of a living being (and I am a member of the set of living beings).

A third type of purpose is an artificial purpose, which is given to a man-made substance. For example, a road sign was erected to help guide travellers, but that purpose was given to it by those who made and use the sign, and does not follow inherently from the materials it is made of.

We should also distinguish instrumental purposes, which are not an end in themselves but a step on the way to some other end. For example, in order to buy a house, I first need to adopt the purpose of saving money, and that purpose will inform my day to day shopping. I am not saving money for the sake of saving money, but because it is a necessary step towards my overall goal of purchasing some property.

All of these are rightly called *purpose* because they are similar in their effects, and they each provide us with a standard of goodness. However, they should be distinguished because they differ in their cause. An artificial purpose arises as an instrumental purpose from an intellectual purpose (since the purpose of the sign-maker was to guide people, and the sign was his means of doing so), an intellectual purpose is the whim of somebody's will, but a natural purpose arises from the underlying inherent and defining tendencies of the type of being in question. It is, consequently, a matter of physics more than anything else; it does not arise from the will of anyone except God, and from God's will only because physics is the description of God's sustaining of the universe. Just as we can study physics without reference to God as its ultimate cause, so we can identify natural purposes without reference to any will, divine or otherwise. Natural purposes are more fundamental to the being than intellectual purposes; therefore while fulfilling a natural purpose is always a good, intellectual purposes may themselves be good or evil depending on whether they are consistent with the natural purposes.

A natural purpose relates to the whole of a being, and not specifically any of its parts, although obviously some parts of the being may be more involved in fulfilling that purpose than others. For example, part of the definition of a living organism is that it is a member of the set of beings with a natural purpose of reproduction. For rational animals such as ourselves, that does not only mean producing the physical bodies of offspring but also ensuring that they reach intellectual maturity; we thus have the natural purpose of conceiving and raising children. Clearly, my genitalia play a crucial role in this, and from this purpose we can determine how sexual expression ought to directed: it is good if it is directed towards a woman, who reciprocates my feelings, and with whom I am joined together in the process of procreating and raising a family (or have the unfulfilled or not yet fulfilled intention of doing so), and an evil otherwise. But the sexual faculty refers to more than just where I put my private parts; it affects my whole body – my sexual desires, my character, how I relate to women in general and my wife in particular, how able I am to protect, help, guide, inspire and look after my family and fulfil my duties as a husband and father. All of these qualities (and more) are needed to fulfil this particular natural purpose; and how good I am as a human being will depend in part on how well they are orientated towards that goal. It is not the fulfilment of the purpose which describes how good I am, but how fit I am to fulfil that purpose. If I fail to achieve the purpose through ill-fortune, then no blame can be attached to me. If I fail to achieve it because of some defect in either my own or my wife's body, then that defect is an evil, but no blame is attached to either of us because we did not choose to have it. However, if I frustrate the purpose though

conscious decision or because I act on a misdirected desire, then I will be held responsible. □

CONJECTURE 14.5.3. *All beings of sufficient complexity have natural purposes.*

Proof:
All beings which can influence others have final causes, since the final cause is just the list of the possible effects of a being. The natural purpose is a generalisation of the concept of final causes to large-scale and complex structures. Whatever the microscopic final causes are, they will combine to give some natural purpose for the complex structure. Therefore all beings of sufficient complexity will have natural purposes, where 'sufficient' here refers to the dividing line where we cease to discuss individual final causes and instead discuss natural purposes. ■

Discussion:
I should also mention the theological motivation for saying that there are natural purposes. If theism is to be accepted, then God created the world not arbitrarily, but with a specific vision in mind. This vision includes various individual beings, each of whom is there for a reason. That reason leads us to the idea that each of them has a purpose in the divine plan, and from that we can determine the natural purposes discussed above. Thus, if theism is to be accepted, then it is inevitable that there will be final causes, and, for the more complex beings, natural purposes. If we suppose that God created things according to a purpose, and without redundancy, then given a form, we can also in principle identify the natural purposes by investigating those individual states of the form where its various unique capabilities coincide. We assume that there is at least one good state, and that good state will be able to satisfy all the natural purposes, so finding a state which can satisfy many potential natural purposes allows is to tentatively identify the natural purposes by investigating that state's capabilities.

If we have natural purposes, then we also have final causes. Thus the existence of God necessarily implies the existence of final causes. Consequently, any form of theism which denies final causality (such as modern liberal theology, based on the enlightenment presuppositions) is thus ultimately self-contradictory and irrational. □

CONJECTURE 14.5.4. *We determine the species to which an object belongs by the formal cause.*

Proof:
Although the final cause is prior to the formal cause in metaphysical priority; the form is prior to the purpose in our sensing and understanding of the object. The reason is that we can directly observe and measure the form; while we have to infer the final cause as best we can. This is particularly true if the final cause is something which happens in the future (a decay of a particle, for example). Since we cannot know the final causes directly, but only indirectly through induction, we identify types of being by their form. ■

CONJECTURE 14.5.5. *Not every natural purpose need be fulfilled by a particular being.*

Proof:
A particle may have several possible decay channels, which we have identified with its final causes. However, in practice, it will only decay into one of these. Therefore some of the possible final causes may be unfulfilled. A similar case can be made for artefacts; although it is designed for several purposes, it need not be used for all those purposes, and on occasion will only be used for one or even none of those purposes before it is destroyed. A car which remains in the showroom until it rusts still remains a car, even though its purpose is never fulfilled. ■

CONJECTURE 14.5.6. *It need not be that every being of a particular substance will be capable of fulfilling all its natural purposes.*

Proof:
The form of a being is defined as a continuous set of Hamiltonians; able to be changed from one to another by infinitesimal changes. Some of these Hamiltonians must be able to fulfil the final cause, as we have discussed. However, the requirement for continuity means that it is not necessarily true that all of them will be able to; it may well be impossible to define a continuous set of Hamiltonians all of which are able to fulfil the final causes. For example, we may consider those beings which naturally grow. A baby is not able to solve quadratic equations, yet it remains within the same form (and thus the same being) as the child a few years later who, after some instruction, can do so. Or this may be due to damage; a car with very slightly damaged spark plug will not, until it is repaired, be able to start; yet because it is barely indistinguishable from a car whose spark plug is slightly less damaged in such a way that it still functions, it remains a car. Thus within the set of Hamiltonians that makes up the form, there is one subset which is able to fulfil the final causes of that form (possibly well or possibly badly); and another (complement) subset which is unable to do so. The subset that can fulfil the final causes must contain at least one member; the complement subset could in principle be empty but it doesn't have to be. ■

CONJECTURE 14.5.7. *The natural purposes of a species can be deduced to a certain extent by considering the formal cause.*

Proof:
This has been discussed above. Given that the natural purpose implies a particular set of forms which have the power to fulfil that purpose, by considering the powers of the possible states of a being it is possible to infer what the purpose could be. Not every power relates to the purpose, and we also consider which powers are accidental to the object; for example as a desktop computer makes a good doorstop; and which are truly essential to its nature. This may be done at least partially by judging whether the object is more sophisticated or complex than necessary to fulfil the particular purpose. However, given that there is some subjective judgement involved, we may only in many cases infer the final causes as probabilities rather than with certainties. Some of these probabilities may, however, be close enough to being a certainty that we can discount the difference. ■

DEFINITION 14.5.8. **Good:**
To be good means to be fit for one's natural purposes; that one's form or action should be

such that the defining tendency is most likely to be fulfilled. △

Discussion:
This definition of good was given above, and is inspired by Aristotle's ethics. It is consistent with our standard usage of the word. The claim is that a moral good is a certain subset of goodness in general. A good car, for example, is one that runs; a bad car is one that doesn't start, or which falls apart while running.

Every being can exist in a number of different states or potentia. Some of these are well suited to fulfil the beings natural purposes; others are less well suited. For example, when a man is suffering from the flu, he is not able to concentrate fully, and therefore not as able to think and reason as clearly as when he lacks the disease. Of course, this is not the only reason why being healthy is better than being sick, but it is still a reason. Thus those potentia where the man is sick are less suited to fulfilling the natural purposes as those potentia where the man is healthy.

Note that there is usually more than one potentia which is best suited to fulfil the natural purposes of the being. There might, for example, be several different ways in which those purposes can be fulfilled, or numerous different states equally suited to fulfilling the purposes in the same way. Goodness is thus not a single state within the form; but refers to a region of states, or a subset of the form.

We should also distinguish between essential and accidental purposes. An essential purpose is one which relates to the species or organism as a whole; it is used to define the species or organism. It is irrational to not seek to fulfil these purposes. However, there might be different ways in which those purposes can be fulfilled. Each of those different ways might imply something more than is just required by the essential purposes. Just as we can make a distinction between essential and accidental properties, we can also distinguish between purposes which are essential to the species or organism, i.e. those which define what it is, and ones which are accidental, i.e. which distinguish the different ways in which the essential purposes can be fulfilled. One good potentia might be related to one set of accidental purposes, and another might have a different set of them, but they both naturally fulfil the essential purposes. Thus we should seek to fulfil one set of accidental purposes (since they are a means to fulfil the essential purposes), but are free to choose which one; none are superior to any other. This is similar to the phenomena of spontaneous symmetry breaking in physics, where there are numerous different ground states with the same minimum energy, but only one of them can be actualised in practice. We thus have a symmetry of the action which is not respected by any actual ground state.

For example, both the philosopher and plumber can equally fulfil man's natural ends. The philosopher seeks to come to as complete an understanding as possible as metaphysics, while the plumber seeks to understand pipes and drains. Neither one of these is a better way of being human than the other. There is, after all, only a limited amount we can learn in a single lifetime, so we are right to specialise according to our own natural talents. The work of a plumber is again defined by a purpose, which is different to that of the philosopher. This purpose defines the good plumber and the good philosopher, and there are different skills which apply to each profession. These purposes are accidental to mankind, but both consistent with the essential purposes, and in terms of those purposes neither profession is superior to the other. Of course, the plumber can decide to retire from plumbing and take up philosophy instead. He would then violate the purposes related to plumbing, but that isn't irrational since they are only accidental rather than essential purposes. He would not cease to pursue the purposes that define humanity; he would merely do so in a different way. □

Objection:
This definition of goodness is based on the purpose or end. But the end of an action is its consequence. Therefore, this simply utilitarianism in another clothing. •

Response:
Utilitarianism judges the merits of an action by its effects; however this definition does not say that an action is good if it fulfils its purpose, since the focus is on *being* rather than *action*. It states that a being is good if it is not only capable but naturally will fulfil its purpose unless prevented from doing so. Utilitarianism judges goodness on account of action, and if the actor is virtuous, that is only seen as good in the sense of being a possible means to the end of a good consequence to the action. Here, however, the capability and inclination of the actor is the end we are interested in. Considering consequences of any actions performed by that agent is merely a means of testing whether or not that end is fulfilled. For example, the utilitarian is content if the footballer scores a goal, and does not care whether that goal is scored because the player is skilful or simply lucky on that occasion. It may be that in a game, the lucky player scores while the skilful one does not. Here, we are judging the footballer on account of his skill, regardless of whether he scores any goals; regardless even of whether or not he plays any games. Of course, we expect the skilful footballer to have a more successful career than the one who relies on luck; but this is only the test of his virtue, rather than the final end in itself.

The one who lives virtuously will most of the time - but not always – satisfy the consequentialist. The one who satisfies the consequentialist will most of the time - but not always – be good according to our measure here. Those exceptions mean that this definition of ethics cannot be reduced to a pure consequentialism. ○

DEFINITION 14.5.9. **Moral good**:
A moral good is that subset of goodness in general which is related to the choices of a rational agent. △

CONJECTURE 14.5.10. *A rational agent ought to be morally good.*

Proof:

(1) The natural purposes are an irremovable part of the definition of a being.
(2) The formal cause is in part determined by the natural purpose. A subset of the form is able to fulfil all of a being's natural purposes.
(3) If a rational agent wishes to be able to fulfil a function, then it ought to be part of the subset of the form which is able and indeed has the tendency to fulfil that function (i.e. it will do so unless prevented by external circumstances). This is impossible if no subset of the form is able to fulfil that function.
(4) If a rational agent wishes to be able to fulfil its natural purposes, then it ought to belong the subset of the form that has the natural tendency to fulfilling all of its natural purposes.
(5) However, the natural purposes define the type of being.
(6) If a rational agent's desires are not in accordance with its natural purposes, then it desires to be something which it is not.
(7) No being can be what it is essentially not.
(8) No if/ought statement where the *if* clause is contradictory is valid.

(9) In any if/ought statement where any alternative to the *if* clause is contradictory, the *if* clause is redundant.
(10) Something which is redundant need not be expressed.
(11) A rational agent ought to belong to that subset of the form that has the natural tendency to fulfil all of its natural purposes.
(12) A rational agent ought to be good.

■

Discussion:
This argument is the key point where I depart from modern thinking, and in particular Hume's is/ought dilemma. Points 1 and 2 were shown above (conjectures 14.5.1 and 14.5.3), while the first part of 3 is close to being a tautology. That it should aim at being in that part of the form which has the tendency towards fulfilling the purpose rather than merely able to do so is also clear; a rational agent can, through effort, overcome the direction the habits of its character leads it, but not for long and not consistently. Eventually either the character has to change or he has to give up on his goal. Far better to have the instincts of character working with you than against you. The man who hates learning might still, through force of will, learn; but he will never do so well as the one with a love of learning. Better would be to foster within himself a love of learning before starting his studies; and then they will be a joy rather than an effort. The second part of 3 is obvious, but important: the form of a being is defined so that a subset of it is able to fulfil all its natural purposes, as discussed above (conjecture 14.5.6). In particular, only a small class of forms are able to fulfil all those natural purposes, and only those natural purposes are fulfilled by only a subset of that particular form. We thus in this way immediately reduce the the number of things a rational agent can consistently wish to be. Note the word 'consistently'; we do, of course, have the capacity for inconsistent thought. A man may wish to be a mosquito, but such a desire is irrational, since he is not a mosquito; and even if, by some miracle of genetic transformation, he could change himself into a mosquito, it would only be at the cost of ceasing to be a man and thus the *ought* clause would be contradictory: it is like saying the part of the subset of men that are mosquitoes. Point 4 follows obviously, and we have shown above that there must be some subset of the form which is able to fulfil all its natural purposes; so the *ought* statement is legitimate.

Points 5 and 6, are, I think, the crucial point where I send Hume into a fit of exasperation. But this follows from the discussion above. Although a being need not be good; nonetheless the being is defined in terms of the good examples of that being. A species is identified in that it is in principle able to fulfil its natural purposes. Now suppose that a rational agent of a certain species does not desire to fulfil the natural purposes related to that species, but instead desires something else. Differences can be expressed in terms of addition, subtraction or change. In this case, the desire must be either contrary to the natural purposes, consistent with their absence, or in addition to the natural purposes of the being. In the case of addition and change, the desire implies a different set of natural purposes; and from that different set of natural purposes we can construct the form of a different type of being[37]. Thus the rational agent in practice desires to be something that it is not. A similar argument can be applied if the rational agent simply wishes to remove one of the natural purposes from its nature. Then the various aspects of the beings form which allow it to fulfil that particular natural purpose would be superfluous. The rational agent's desire would be just as well served by being something that lacked those aspects of its form;

[37] I am excluding the possibility here that these different set of natural purposes is available in a subset somewhere else within the form. If so, the differences in desire are accidental rather than essential to the nature of the being; I am here focusing on the desire to be different than what it is essentially rather than accidentally (see point 7). For example, a teacher may legitimately wish to be an astronaut, or an astronaut a teacher, as these differences are accidental to the nature of man. However, a teacher may not legitimately desire to be a spaceship, as a spaceship is essentially different from a man.

and given that we earlier stated that we should not describe those putative purposes where a being is more sophisticated than it need be to fulfil them as natural purposes; in practice the rational agent would be desiring to be such a being. The rational agent would again desire to be something it is not; an impossibility. Therefore the only rationally legitimate desires of a being are those in accordance with its natural purposes. The other possibility is that an agent may desire to retain its same form while having a different set of natural purposes; however, given the relationship between the natural purpose and the form, this is incoherent.

The rest of the argument is straightforward. If a being desires to be something it essentially isn't, then that is contradictory, and one of the rules of the if/ought statement is that the if clause must be possible. If the only alternatives to the *if* clause are contradictory, then the *if* clause is redundant and need not be stated. We then, utilising the definition of goodness (14.5.8) in the last step to reach the desired conclusion.

The conclusion is in accordance with our common expectations. The defining purpose of a teacher is to teach, therefore a teacher ought to teach. A soldier ought to fight when the battle is called. A cricketer ought to try to score runs. These statements are not controversial; they follow from this line of argument. □

CONJECTURE 14.5.11. *Rational agents ought to seek goodness in themselves.*

Proof:
Rational agents ought to be good, as proved above. One aspect of goodness is to have the will directed rightly, that is, towards the good. A will that does not seek the good of the being which it is contained within is not directed rightly; thus this type of will is vicious. The corollary of this statement is that a virtuous will is one that seeks the goodness of the being within which it belongs to. As rational agents have wills (definitions 4.4.1 and 4.4.9), this means that rational agents ought to seek goodness in themselves. ■

CONJECTURE 14.5.12. *Rational agents ought to seek goodness in others.*

Proof:
The argument is similar to the one above.

(1) If a rational agent desires that another object be in accordance with its natural purposes, then he ought to desire that it belongs to the subset of its form that has the natural tendency to be able to fulfil its natural purposes.
(2) If a rational agent desires that another object be in accordance with its natural purposes, then he ought to desire that it is good.
(3) No object can be what it is essentially not.
(4) The natural purposes of an object define the type of being that it is.
(5) If a rational agent desires that another object not be in accordance with its natural purposes, then he desires that it be something which it is essentially not.
(6) A redundant *if* clause in an if/ought statement need not be stated.
(7) A rational agent ought to desire that other beings become good.

■

Discussion:
Point 1 is our standard starting point; this is again close to a tautology and thus sufficiently obvious that it can be accepted as a premise. Point 2 follows from the definition of goodness

(definition 14.5.8). Points 3, 4 and 5 parallel the discussion in conjecture 14.5.10. The conclusion follows immediately. □

CONJECTURE 14.5.13. *For certain natural purposes, there may be degrees of goodness.*

Proof:
This follows from common experience. Wayne Rooney is a good footballer; Pele was a better one. Thus there are degrees of goodness if the natural purpose is footballing ability. To say that something is able to fulfil its natural purposes is to say that there is a certain non-zero probability that it will in practice fulfil them given a particular situation. Clearly, some objects may well have a greater probability than others of being able to fulfil those causes in practice. By saying that something is better, we are saying that it has a greater probability of fulfilling its causes if it follows its inclinations and instincts. But saying something is better than another is only possible if there are degrees of goodness. Therefore there are degrees of goodness associated with at least some (though not necessarily all) natural purposes. ■

CONJECTURE 14.5.14. *Rational agents ought to seek the highest good for an object consistent with the nature of the object.*

Proof:
We have already shown that goodness is to be preferred over badness. Yet the difference between goodness and badness is one of degree rather than kind. Thus we have already shown that in one sense that which is better is too be preferred over that which is worse. However, this is not the same as saying that that which is best is to be preferred most of all. If the formal cause is defined through the final cause, then equally we can consider that within the possible arrangements of matter that constitute the form, there will be at least one which best allows the object to fulfil its final cause; that is to say that it has a certain large probability that the natural purpose will be achieved given the nature of the object. We may consider that there are other forms which have a lower probability of fulfilling the final cause. The reasoning now continues in the same way as before. If we are satisfied with something that is merely good (a probability p_1 that the natural purpose will be achieved), rather than the best possible (a probability $p_2 > p_1$), then we are effectively desiring that the object is something which need not, given its nature, reach the level of p_2 but only p_1. We thus desire that the object is something other than what it is, in this case something less than what it is. This is impossible. We thus conclude that the original premise, that we ought to be satisfied with something less than the best, is false. Instead, we should only be satisfied if the object achieves the highest possible good consistent with the nature of that object. ■

CONJECTURE 14.5.15. *Only something which is essentially good can make something that is not essentially good necessarily good (either directly or through an instrument).*

Proof:
There is a region in each form which is capable of fulfilling its natural purposes, which is the the definition of what makes a being good. The form of a being may extend beyond this

region (which is what we have been considering) or it may be restricted to this region. By essentially good I mean something whose form is restricted to the region where it is good. By necessarily good I mean something that remains good permanently. Now if something is not essentially good, then it can stray from the region in its form where it is good, and unless constrained to that region by some internal or external power then it will eventually stray from the region of goodness (since everything that can happen will eventually happen: this is the definition of 'can'). However, it cannot be constrained by its own nature, because if that were the case it would be essentially good (since its nature would be to keep itself in a state of goodness). On the other hand, unless it is constrained by something good at performing that action, then that constraining force will be insufficient (since if the force is not good, then it will eventually fail, as follows from the definition of goodness). To maintain that constraining force permanently, that being doing the constraining needs to be necessarily good in that respect. So only something necessarily good in this respect can perform the constraining. But if it is not essentially good in that respect, then something cannot constrain itself to be perfectly good, as already shown. If it is constrained by something else, then that thing must either be constrained to be necessarily good, or it is essentially good. This leads to a series of constraining powers, which can only terminate with something that is good essentially in its ability to constrain (the series cannot continue indefinitely because it is a series of instrumental causes). This chain must terminate with God, since the power to constrain something in a state of goodness depends on the existence of the constrainer, so if the constraining power is essential to the being, then so must its existence be essential to it. But only God exists essentially. Therefore God can be the only being which is essentially good in this respect. But God is essentially good in every respect (see conjecture 14.5.16 below). Thus only something which is essentially good in every respect can constrain something which is not essentially good to be perfectly good. Since the act of constraining into a good state is part of making something perfectly good, this means that only God can make a being become perfectly good. ■

CONJECTURE 14.5.16. *God is essentially good.*

Proof:

(1) To be good means to be in accordance with one's natural purposes.
(2) Every being's form describes at least one state which is good. It may also describe states where, if one of these is actualised, the being would not be good.
(3) A being which is not good must have have a form which covers various different possible states; at least one of which is able to fulfil its natural purposes, and at least one of which is not able to do so.
(4) God is simple; pure actuality with no potentiality.
(5) Therefore God must have only one possible structure described by his form.
(6) This possible structure must be able to fulfil its natural purposes.
(7) Therefore God is essentially good.

■

Discussion:
The first point is the definition of goodness (definition 14.5.8). However, there is a complication. So far, my discussion of natural purposes has been restricted to physical beings. The natural purposes are an expression of the macroscopic effects of the first sense of final causality discussed in definition 4.5.10, and this sense only applies to material beings. However, God is immaterial, so definition 14.5.8, as precisely stated, can't apply to God.

It therefore seems as though the argument is invalid, since points 1-3 refer to a physical being, and in 4-7 I apply this to an immaterial being. My way of evading this objection is to extend the definition of goodness so that we can use the word in an analogous sense to apply to God.

I originally defined God as the termination of every chain of efficient causes (13.1.12), with God's will playing the role of final causality, linking Him with the various effects. Here I am arguing that every species can be defined in terms of a set of natural purposes. I define God as an efficient cause, while my discussion in this chapter relies on final causality. For this proof to be valid, I need an alternative definition of God, expressed in terms of either natural purposes or the power to bring about certain effects (as natural purpose is merely a macroscopic view of the beings final causes, and each final cause implies a power to bring about an effect). This alternative definition needs to be equivalent to definition 13.1.12, in part because later in the proof I rely on God being pure actuality, which is derived from definition 13.1.12, and in part because I would rather avoid fallacy of equivocation as I build on this conjecture. So I need to uniquely define God in terms of a list of effects which God and only God can bring about, and show that this definition is equivalent to definition 13.1.12. We know that for God this list of effects is unbounded: this is a statement of God's omnipotence. So just as we can define mankind as a particular one of the possible species which can fulfil the various biological and intellectual ends, so can we define X as the being which can fulfil unlimited ends in terms of the generation, corruption and motion of other beings, or the actualisation of the potentia of other beings. As shown in 13.3.9, only God can fulfil the ends ascribed to X. In this way, the definition of goodness can be extended to cover God.

Nonetheless, this subtle difference in definition means that the word *goodness* as applied to God is only analogous to how we would apply the word to (for example) a man. It is right to use the same word, since the sense of goodness discussed in this paragraph is to God what definition 14.5.8 is to a man; but the two senses of the word are not identical because God is fundamentally different to any material being. It should be recognised that whenever I speak of the goodness of God, I mean this analogous sense of the word.

The second point follows from self-consistency: if a being is determined by its natural purposes, then it must have at least one state which is able to fulfil those purposes. This is true both for the definition of goodness as applied to material beings, and the analogous definition applicable to God. Given that every being has a natural purpose (even if that is merely described by some equilibrium state), this demonstrates the second point of the proof. We construct the form of a being in terms of a set of continuous stable states, one of which is this good state. The form might just cover the good state, or it might extend beyond it: while the discussion remains at a general level, we have to keep both options alive. If a being is not good, then there must be more to the form than merely that region which is able to fulfil its natural ends. That God is simple was shown earlier (conjecture 13.2.7). This means that God's form only has one possible 'state' (the word state is not really appropriate for God; here I use it analogously). Given what was said above, this can only be that point able to fulfil God's natural ends. Thus God is not only good, He cannot be anything other than good as this would contradict God's simplicity.

Because God is essentially good, God's goodness is interchangeable with the concept of God's being. However, for us, goodness (if achieved) is merely an accidental property, and therefore our goodness is not interchangeable with being. This is another way in which the concept of the goodness of God is only analogous to the sense used when we discuss a good man (if any such men exist). However, the concepts are close enough that we are not being deceptive by using the same word to describe them. □

Discussion:
This means that God is capable of maintaining a being in a state of perfect goodness. Since

God is the only being of pure actuality, and only the being of pure actuality can be perfectly good (since everything else is created by God, and thus can in principle lack the good of existence; existence is a good because without it a being is not able to achieve anything, including its natural purposes), only God can maintain a being in a state of perfect goodness. □

DEFINITION 14.5.17. **Disorder**:
A rational agent is disordered if it deviates from the good, though not through a choice but because of disease, damage, or circumstance. △

DEFINITION 14.5.18. **Defect**:
A defect is a deviation from the good due to the choices of a rational agent. △

CONJECTURE 14.5.19. *Beings may differ from the good either from disorder or defect. However, while a rational agent is always morally culpable for their defects (created by themselves), it is not necessarily culpable for their disorders.*

Discussion:
This conjecture is, I think, relatively obvious, but also important. A man who loses his limb by accident or disease, or because of the someone else's evil, should be pitied and helped; a man who freely chops it off (other than for a medical or other emergency) should be treated as insane. A man who suffers from some strong temptation to do evil is not to be criticised, since he cannot help but have the thoughts come to him; but the man who gives in to the temptation and is unable to control his actions, or who comes to regard what he is being tempted to do as good or less evil than it is on account of it, is to be strongly condemned.

I should mention that there is a difference between controlling a desire and suppressing a desire. I think that we all have temptations within us which we find it hard to overcome (unless we have succumbed so much that we do not see them as evil). For some people these are fairly small, almost trivial matters; for others they reflect a large part of their lives. Such temptations are rarely of our choosing; they represent a disorder with us. But to give in to them, with all the consequent harm that does for us, would be a defect and morally wrong. To suppress a desire is to try to stop yourself from feeling it; to control a desire is to have the internal strength to be able to laugh at it when it comes, and then continue on your way. While to not feel the desire in the first place would be best if practical, most attempts to suppress it are counter-productive. It is better to just recognize it, both as a flaw you desire to be without and as something you may have to live with, but learn to control it. It means regretting your failures, but not allowing yourself to be consumed by them. Sometimes, after enough time, the harmful desire will disperse and be replaced by a healthy desire; on other occasions it remains throughout your life. But living with such a temptation is better than surrendering to it: that will only feed the monster, cause you to identify yourself with it, and will cost you your humanity and ultimately your happiness. On the contrary, the one who overcomes the greatest challenge receives the greatest reward. However, if you give in, there is little human hope. The more one becomes accustomed to evil, the harder it becomes to resist or even recognize it as the destructive force in your life that it is.

As shown above, we are not strong enough to permanently control temptation without God's help (either directly or through a series of instruments); at least, though we may win some battles, maybe even most battles, we will not win them all. Unfortunately, the demand is that we win them all. God's purpose in creation is in agreement with the natural purposes of each type of being, since God is the creator and designer of those beings; and the natural purposes of the being are the foundation of the design of the form. Thus for a being to be good is to be aligned with God's purposes for that being, and God desires that things should be good. Therefore God's purpose in creating mankind is to create a perfect society of men and women as they ought to be. Without injustice, without pain, without death. This is impossible unless the people within it are perfectly good, which means that they can have neither defect nor disorder. If there is any imperfection (i.e. a non-zero probability of doing harm), no matter how small, eventually it would be realised, and once realised, magnified. This is not possible unless the people are continually renewed and strengthened by God; but God only does so if the people desire for Him to do so, since He would otherwise be overriding their will, and thus destroying a key element of their humanity. Thus only those who desire to be good, and who desire that God should continually make them good, and accept His method of doing so, are fit for this perfect society. If anyone flawed (as everyone today is flawed in at least some respects) enters the perfect society, then it will immediately cease to be perfect. For defects, this is impossible, since a defect originates in a desire aligned against a good. For a disorder, it is not necessarily impossible, since someone may have a disorder against their will. However, if somebody does not accept that only God can remove their disorder and trust Him to do so, or does not wish that this disorder is removed from them, then the disorder will also prevent them from entering the ideal society that God seeks to prepare. □

CONJECTURE 14.5.20. *To deliberately use some faculty in such a way that is contrary to its natural end is an evil.*

Proof:
A complex being, such as ourselves, has many parts, and some of these parts or faculties are particularly directed towards some of the natural ends of the being. For example, the intellect is that part of us which is directed towards the acquisition of true knowledge, while the spleen has other natural ends. Thus many parts of beings such as ourselves have an individual natural end; and the proper use and functioning of these parts is required if the being is to achieve its natural ends. If we use one part of a being in a way that is contrary to its natural ends, then the being taken as a whole will be acting in a way contrary to its natural ends. However, goodness of action is defined as acts which are consistent with the natural ends for the type of being which is acting. Thus to use some faculty in a way contrary to its natural ends is inimical to goodness, and is thus an evil.

This refers to positive use of a faculty in a way that is opposed to its natural end. For example, the purpose of our teeth is primarily to assist in our eating process. However, there are other uses of teeth (for example strumming against them with a finger) which are not immoral, since they do not hinder using the teeth for eating, nor are an imitation or perversion of the eating purpose (for example by chewing good and needed – by someone who might otherwise receive it – food and then deliberately spitting it out, performed just for the pleasure of chewing rather than for any nutritive benefit). ■

CONJECTURE 14.5.21. *To deliberately use some faculty in such a way that bears no relation to its proper use need not be an evil.*

Proof:
Although our faculties are directed towards various goals, some of them need not fulfil those goals continually. For example, the intestines need not continually digest food in order to achieve its purpose towards the end of keeping us alive and healthy. As the intestines fulfil their end, by extracting the nutrients from the food and dispersing them into the body, then that is a good. If they were to act contrary to that end (for example by sucking nutrients out of our body and into our excrement), then that would be an evil. However, while the intestines are not digesting food, then there are various movements they can make which are irrelevant to digestion: neither absorbing nutrients, nor removing them. These movements (as long as they do not injure the intestine) are ethically neutral; neither acting towards a natural end nor contrary to it. An ethically neutral action is not an evil. Thus (all other things remaining unchanged) it is not evil for our intestines to move in such a way that is unrelated to their natural function.

Let us suppose that the sake of this example that the natural end of our hands is to manipulate external objects, while the natural ends of our feet are for walking. It is, however, possible to walk on one's hands. This is a use of that part of ourselves for something other than its natural purpose. However, this need not be an evil. An evil is a privation or lack of goodness. By walking on our hands on those occasions when they are not required to manipulate external objects, we are not harming their ability to do so at those times when they are so required (unless some injury results, which would be an evil). Thus we are not preventing our hands from achieving their natural ends, or having them act in a way contrary to their natural ends. In this way, we are not doing something harmful, and this is morally neutral.

Evil is a privation of good. To commit an evil is to do something that is opposed to goodness. There are three types of action: those that increase goodness (are in accordance with natural ends); those that reduce it (are in discordance with natural ends); and those that have no bearing one way or another. But a morally neutral action is not necessarily an evil; it does not involve corrupting something from a state of goodness, but instead maintains it at the same level of goodness. In the same way, we can use some faculty in a way that is in accordance with its natural ends, or contrary to its natural ends, or neither in accordance nor contrary to its natural ends. The last of these, which includes examples as walking on ones hands, or using one's knees to act as a table, is analogous to the morally neutral action. Thus this need not be an evil. It might, however, be an evil, if it prevents that or some other faculty from being used to fulfil its natural end at the time when it needs to be used, or if it perverts the proper use of the faculty.

In many cases, we not only need to use the faculties correctly, but also apply them in the right situation for them to achieve their ends. We might need other substances to achieve the end, and these would need to be of the right form; for example to nourish ourselves we should eat bread rather than concrete. To direct the digestive faculty towards concrete is a misuse of it, even if we correctly chew it, swallow it, and pass it into the stomach. Thus to attempt to consume concrete, even if it does not harm our digestive system or prevent us from consuming genuine food at another time, would still be an evil because it mocks the proper use of the faculty. The error is in the improper direction of the digestive faculty. Our digestive faculty is naturally directed towards food; this follows from both its natural purpose and the nature and forms of food. To instead direct it towards concrete is to make one of two errors: to say that our digestive system ought to be directed towards concrete in addition to food, or to say that concrete ought to be food. However, both of these are against reason, as they are trying to say that either our digestive system our concrete ought to be something they are not. This is irrational, and consequently an evil. However, merely

placing concrete in one's mouth without treating it as though it were food (by attempting to chew it, swallow it, and so on) does not make this error; it might not be wise to do so, but it is not an evil in the same way.

If, on the other hand, desiring to walk on our hands, we disfigure our hands so that they are better suited to it (perhaps by making them more closely resemble a foot), then that is a moral evil, since it hinders them from fulfilling their natural end on other occasions.

A corollary is that non-use of a faculty is not an evil unless it frustrates the body from achieving its natural ends. To fast for a short period is not an evil; to fast in excess and thus to starve oneself is certainly an evil since it frustrates the natural end of continuing as a living organism. ■

Objection:
It is possible to use an object in a way that frustrates its natural purpose, but still leads to a good. For example, one can and ought to tear up a shirt to bind a wound if there are no better bandages available. •

Response:
This is a case of balancing different evils. It is an evil against the shirt for it to be torn up and used to bind a wound. However, it is also an evil for the person to have an open wound (leading, for example, to the possibility of infection and long term impairment to health). Using the shirt to bind a wound is an evil against the shirt. Not using the shirt is an evil against the person. Clearly the person is more significant than the shirt, firstly because the shirt is just an artefact (whose purposes are for the sake of mankind) while the man is a natural substance, but also because a person has considerably more natural purposes than the shirt (for example a person is capable of perusing intellectual ends, while a shirt cannot). The infection or blood loss could lead to serious health problems (and possibly death) for the person. Thus the desire to promote the person's goodness ought to be stronger than our concern for the shirt. This is regardless of the outcomes of the action. It may be that it is not necessary to bind the wound, so we might be damaging the shirt with no benefit to the person. But there remains a risk of infection, our desire ought to be to reduce that danger. The actions should follow from this desire, and one should bind the wound. Obviously, if we had a bandage available, then it would be wrong to use the shirt, in part because we would be destroying the shirt when we have no need to (and in part because the specialised bandage would probably be more effective at binding the wound than the improvised shirt). Our desire is to help the person, not to destroy the shirt. The destruction of the shirt is an unavoidable consequence of the actions motivated by the desire to prevent evil to the person.

In every example of this sort, we need to make a judgement. Sometimes the judgement is easy; other times it is a much closer decision. In those circumstances when doing an evil to A is the only way to prevent a possible evil to B, we have to decide what to do based on how serious the two evils are and how likely they are to occur if we do or don't do the action. Some evil is inevitable: our intention in these circumstances should be to minimise it. ○

DEFINITION 14.5.22. **Man:**
A Human being is defined as a species of living organism with a genetic code within a certain area of the genetic landscape, with the (species-level) form of Man containing the individuals known as Socrates and Marie Curie. △

CONJECTURE 14.5.23. *Men are inherently living, rational and social animals. Their natural purposes therefore incorporate the natural purposes of living beings; animal beings; rational beings and social beings.*

Proof:
In principle, we should derive this from the form of man. However, it is easier to demonstrate this through direct and obvious observation. ■

Discussion:
So, for example, as living beings we have the responsibility to try to prevent or cure ill-health in ourselves or others, and do all we can to prevent others from dying. We also obtain from this the natural purpose of reproduction and growth, and combined with details from our animal nature and particular biology we give rise to the purposes from which we can derive reproductive and family ethics. Our intellectual nature arises from the natural purpose of the acquisition of knowledge, from which we gain the moral imperative to continually educate both others and ourselves. A society is a community of people working together for the good of all (with goodness as established by the purposes derived from our living, intellectual and animal nature); and this is related to the purpose that, as far as we are physically able[38], we should contribute to society to the extent that we receive from it. Structures in society naturally rise to fulfil various purposes, and from those purposes we derive property, political, and corporate ethics, and so on; with each corporation or governmental structure arising from the necessity of providing a particular service which the wider society needs; from that purpose we can define the standards of goodness appropriate to the organisation; and also the duties and special virtues of those working in those organisations. □

DEFINITION 14.5.24. **Life:**
A living being is that which has the natural purposes of growth, the absorption of nutrients, reproduction, and maintenance and repair of its complex structure against disease and damage, and consequently the capacity to perform these functions. △

Discussion:
This definition is reasonably standard, although some may add more powers for a living organism. It should be particularly noted that life is defined in terms of its natural purposes (or tendencies); the primary difference between a living organism and something non-living are these powers. Thus the natural purposes of every living organism include growth, the absorption of an appropriate amount of nutrition (and consequent use of it and the emission of waste products), the ability to sustain itself, and reproduction. The organs of a living body will mostly be directed towards these or other natural purposes.

For example, we can see that one of the natural purposes of a living organism is to continue living. Thus to inhibit an organism's life is an evil, and to deliberately do so a moral evil. From this, we conclude that it is morally wrong to deliberately kill any life, except when we have no choice (for example, if the alternative would be worse; this means that it is not wrong if, as a last resort, we end a life in defence or for food). This may be through deliberate damage or injury, or by preventing the organism gain the sustenance it needs. It is no less evil to harm ourselves, through consuming far less than we need,

[38]We should help the disabled or infirm achieve as much as possible with their lives, but recognize that they are not in a position to contribute as others would - an evil, but not a moral evil because it is not by choice, but rather the evil resides in the damage done to the person by the disability.

consuming far more than we need, lack of exercise, over exercise to the point of exhaustion or injury, and so on: we have a responsibility to look after our bodies (although, again, there may be exceptions when it may be right to fail to keep this responsibility if it is the only means by which we can fulfil a higher purpose).

We are called to ensure that, as much as it is possible, everyone has access to what they need to survive; to feed the hungry, give drink to the thirsty, clothe the naked, bring heat to the freezing, teach people to farm efficiently and manage and budget their resources, provide healthcare, and so on. These things are moral imperatives; how they should be efficiently accomplished, however, is a much harder question of economics, and I don't think that anyone has yet found a good solution (though numerous bad solutions have been proposed). The problem is made harder by a lack of moral awareness in society, and that too many of us focus on distributing according to human wants first of all rather than seeing human needs as the priority (if there is enough left over after providing for our needs, then we can and should discuss how to distribute it between the infinite human wants).

Now, I should make one caveat: while this applies to all life, and we should not harm anything unless it is necessary, there is a hierarchy in life, according to the various powers of the types of life. Thus viruses and bacteria, which have the least powers, are least valuable; then plants and fungi; then animals which also have the powers of sensation and self-motion, and some degree of rationality come next, with the fully rational animals (ourselves) at the top. The reason for this is that all the powers of a living organism depend on its life as a first principle. If one kills a bacteria, then one just ends its life and nothing more. To kill a man, however, not only ends his life but also his powers to sense the world around him, and to understand what he is sensing. There is, then, no evil when a man's immune system destroys the bacteria invading him, since otherwise they would weaken him and his life is more valuable than theirs. It is evil to destroy bacteria for sport (because one can), but not for protection. □

Objection:
This is an attempt to derive a beings purpose from its form. This is just the naturalistic fallacy, and should be objected. •

Response:
This is precisely what we are not doing. We have defined life in terms of various functions and purposes. We have then (implicitly) stated what forms might satisfy those purposes; in particular we have noted that any being whose form includes at least one state which satisfies all those purposes can be thought of as living. We then construct the forms of various observed beings and species, and noted whether this condition is satisfied. We have not gone from form to purpose; we have started with tendency and then constructed the possible forms which could fulfil that purpose, then compared these against the forms we observe in nature. ○

Objection:
One of the definitions of the final cause is the stable or meta-stable state to which a physical process tends. Yet every living creature inevitably dies. Therefore the natural purpose of life is death; and therefore, if the previous argument is accurate, the best thing you can do for a living creature is to kill it. •

Response:
The final cause is the equilibrium state if external conditions remain constant. The natural purpose of a living organism is that which it tends to if it does not have external interference. Most death, on the other hand, at least for the higher organisms such as animals and plants, is caused by decay not from the organism in itself but from external factors. Indeed, in the more advanced organisms, the body does all it can to fight off these invaders. This is most obvious if the death is caused by damage (such as falling off a cliff, or beheading), but also if it is caused by contracting a disease. Other forms of death are more subtlety caused by external effects; for example genetic degradation from radiation or excessive sunlight leading to cancer, or are caused by a failure to properly look after the body (for example, a heart attack or stroke caused by high cholesterol), or by some portion of the body no longer functioning correctly, perhaps because of some congenital flaw or environmental damage. Occasionally, some part of the body might just wear out through overuse. Our bodies are under constant attack, and eventually we succumb. In all of these, it is some flaw in the body, damage, or external attack which causes it to fail to fulfil its final cause, which is to remain alive. Recall that it is not true that every being necessarily fulfils its natural purposes; rather only those which are good can do so. The fact that every living organism eventually dies may mean that no living organism is able to sustain itself in a good state perpetually.

We thus have two competing explanations for why a living organism may die: firstly, that the final cause of a living organism is death, and that those creatures which are healthy are bad; secondly that the natural purpose of a living organism is life, and that those creatures which are unhealthy are bad or have become bad under external pressure or natural degradation[39]. In this second case, the fact that all living organisms eventually become unhealthy does not affect the point; in our discussion of goodness above we did not require that any actual being (as opposed to the potentialities described within the form) should be good, and even less that it should remain so perpetually. As long as there is a possible cat which is able to survive until the next moment, then the natural purpose of a cat might be to remain alive, and this does not contradict anything above. However, if the natural purpose of a living cat is a dead cat, then why does it need the capacity to remain alive even for a short moment? Such a capacity would be superfluous to the nature of a cat, and we reject superfluous purposes as possible natural purposes.

It may be further objected that with the set of states described by the form, there is also a cat which is on the point of death. However, being able to sustain oneself is only one of the distinguishing powers of life. A form is valid if there is a point within the states contained within it which satisfies all the natural purposes of the being. A dead cat cannot reproduce, absorb nutrition, or grow; a living cat can. Therefore death cannot be the natural purpose of a living organism; it is instead life. This is, of course, in accordance with our intuition. ○

Objection:
It is said above that it is immoral to kill anything except when necessary, and the higher life forms most of all. Yet we raise and kill animals to eat them; and this is commended in the bible. Jesus himself ate meat. Therefore such a doctrine is incompatible with Christianity. •

Response:
I believe that there is some truth to this line of thought, and I myself am vegetarian largely in response to more developed forms of this argument. However, the argument does have a flaw: we are entitled to kill animals if it is necessary for our survival or even our good health. For most of human history, a good diet required eating meat, as there were certain

[39] I am not using 'bad' in a moral sense here, but its more general sense.

nutrients which are difficult to obtain without meat. Even today, when vegetarianism is a viable option (and perhaps even healthier and better than being an omnivore), it is not so everywhere and it is not always easy. Of course, even a vegetarian kills plants in order to eat.

The biblical position is somewhat more involved. The claim is that initially people were only permitted to eat fruit (and possibly beans, nuts and so on), although this instruction might not have always been obeyed.

> And God said, "Behold, I have given you every plant yielding seed that is on the face of all the earth, and every tree with seed in its fruit. You shall have them for food. And to every beast of the earth and to every bird of the heavens and to everything that creeps on the earth, everything that has the breath of life, I have given every green plant for food." And it was so[40].
>
> And the Lord God commanded the man, saying, "You may surely eat of every tree of the garden, but of the tree of the knowledge of good and evil you shall not eat, for in the day that you eat of it you shall surely die[41]."

Both these quotations refer to the time before evil enters the world. Thus, if these stories are to be taken seriously, the initial state, or design of man and animals would be to eat only plants as necessary, in accordance with the implications above. This changes a little later in the story,

> And God blessed Noah and his sons and said to them, "Be fruitful and multiply and fill the earth. The fear of you and the dread of you shall be upon every beast of the earth and upon every bird of the heavens, upon everything that creeps on the ground and all the fish of the sea. Into your hand they are delivered. Every moving thing that lives shall be food for you. And as I gave you the green plants, I give you everything. But you shall not eat flesh with its life, that is, its blood[42]."

A commandment may change either because circumstances change (i.e. the fruit required for good nutrition was no longer readily available), or as a concession to human weakness or stubbornness[43]. If the reason is the first of these, then we might say that given that circumstances have changed again, the consumption of meat should now be avoided for the Christian, although it is unclear that circumstances have changed enough. If it was a concession, then it might well still apply. I do not therefore believe that a Christian should be condemned for killing to eat meat[44]. ○

CONJECTURE 14.5.25. *Life is the most important of the purposes which a living being can have. To destroy an innocent being's life is the greatest evil one can do to it.*

Proof:
I wish to explore two reasons for supporting this. The first is that (almost) all other goods

[40]Genesis 1:29-30.

[41]Genesis 2:16-17.

[42]Genesis 9:1-4.

[43]c.f. Matthew 19:8.

[44]See also the discussion in Romans 14; there is also an important issue in the way the shedding of blood was an important aspect of the doctrine of atonement in the old covenant, with the new covenant doctrine of atonement different from but still derived from that. In this way, the eating of meat played an important role for Christians in redemptive history as a symbol of and substitute for our own deaths, both physical and spiritual.

depend on life. Thus if one terminates life, then one also removes the opportunity for all other goods. Secondly, of the moral choices we can make, the decision to end a life (or cause serious injury) is one of the few that is irreversible. Most other harm to either the body, mind, character, reputation or possessions can in principle be reversed or redeemed; but it is not in our power to bring back the dead. It is no use (in this sense) regretting murder; once done the deed cannot be undone. Of course, this does not deny to God the power of resurrection; indeed the Christian believes that all will be resurrected to face judgement[45]. However, we are responsible for our choices and character, and we ought not do or be evil even if good later comes from it by some means outside our control. ■

Discussion:
There is a caveat I should make here. I believe that taking a life (or otherwise causing great harm) is always an evil. However, it is not the only evil, and sometimes, when we have to choose between evils, it might be the right choice. I wish to mention two occasions when this might be the case. The first is in defence, whether self-defence or defence of another. This case seems to be clear; we are left in a position where we either have to disable the aggressor, perhaps even to death, or face another death. In this case, since some injury is inevitable, justice demands that it is better that the aggressor is injured than the innocent party. Of course, such actions should ideally only be taken as a last resort and with the minimum amount of damage inflicted (although I think that in such circumstances the benefit of the doubt should be given to the defender). The Christian concept of the just war is the extension of this principle to a larger scale.

The second occasion is perhaps a little more controversial; and that is when the principle of not taking a life conflicts with the principle of justice. Justice is the principle that everyone should be treated as they deserve, most especially according to their deeds. It is unjust to give harm to the saint; it is equally unjust to benefit the robber as we do the saint. It is also unjust to give a greater benefit or lesser punishment than is deserved, or a harsher punishment or smaller benefit. Many people have held that there are some deeds which deserve death, for example the deliberate taking of an innocent life, with the punishment proportional to the severity of the crime. If this is the case, to give the criminal a lesser sentence would be to violate the principle of justice. To give a lesser punishment would, it is claimed, devalue the life of the one who was killed by the transgressor. It would say that a fine or imprisonment would be sufficient to repay the damage done to the victim. But if the victim's life is of inestimable value, then that cannot be the case.

It is not my intention to say here whether this is indeed the case: whether there is any deed that deserves death, or to express my own opinion; I merely wish to point out that it might be so (Feser and Bessette, 2017; Schieber et al., 2013).

I do not believe that there are any occasions other than these two which merit overriding the good of life. All goods are either goods for society in general or for individuals. No good for an individual is more important than life, since there are always alternatives to killing a person that relieve the pressure without causing the evil or causing a lesser evil. Society, on the other hand, is just the community of people, bound together by inter-dependence, including the principle of justice. Justice is, however, the only societal good that may involve restricting goods to members of the society; otherwise one builds up society by building up either the individuals within it or the bonds that connect them. To take a life diminishes society. □

[45]Hebrews 9:27, Revelation 20:5

Objection:
While ending most lives is indeed wrong, for some people to live is merely to prolong suffering, or to bring into life is to bring into a time of predetermined misery. In these cases, to die is better than to live. Therefore to take a life in these circumstances is good. •

Response:
Suffering, if it is to be regarded as an evil, is merely the privation (or lack) of some good, or a symptom caused by the lack of some good (for example proper health). Evil does not exist in itself; rather it is the absence of goodness. To end a life is not only to deprive the being of that good, but of every good. Therefore to end a life is morally worse than allowing someone to live on in suffering.

Of course, this does not mean that we should see pain as a good; on the contrary we should do all else we can to relieve it, but not by sacrificing other goods to end it. ○

Discussion:
I have so far spoken of self-sustaining nature and nutrition as natural purposes of living organisms. The capacity for physical growth does not affect a great deal morally, but reproduction is of significant importance. The natural purpose of reproduction (and the consequent overseeing of the child's development) is at the centre of the natural law treatment of sexual morality. While a full discussion of sexual morality goes beyond the scope of this work, I will say a little. Only those sexual desires which are consistent with the natural purpose of reproduction and raising children are good; a morally good sexual action is one where one does not deliberately hinder this purpose. So, for example, intercourse with an infertile woman is an evil, but not a moral evil, since the evil is in the disease, damage, or age that made her infertile rather than the intention of the agents. Sexual activity by a man with another man is always an evil, because the natural purpose is hindered not because of a disorder in the nature of one of the parties, but because they are of the wrong nature. Homosexual desire, if unchosen, is not a moral evil – though it obviously is an evil, because it focuses sexual desire towards something other than the right substance needed to fulfil the natural purpose of reproduction which gives rise to the desire; and the people who experience that desire on average no more or less evil than anyone else, but to deliberately act on that desire (or to make that desire part of one's self-identity) is a moral evil. To encourage such behaviour is itself an evil and an act of hatred towards those who practice it, since hatred is a deficiency in the desire for good. The only loving response to those with same sex attraction is to discourage homosexual activity, the harm and damage that follow from it, and the evil that it reinforces. □

CONJECTURE 14.5.26. *The natural purposes of animal life include self-motion, sexual reproduction, and the ability to sense, process and react to external stimuli such as light and sound.*

Proof:
This can be treated as a definition of animal life; all those we classify as animals satisfy these purposes. ■

Discussion:
Of course, almost all living things react to stimulus; animal life however differs in the way in which it absorbs and processes the information, through sensory organs and the nervous system. Otherwise, the moral implications of these powers are limited. For example, our ability to move implies that to restrict someone's mobility (for example by denying access to some public area) is generally not good (although there may be certain specific reasons to

do so in certain cases, where since the world is not ideal one has to weigh which is the lesser of various competing evils, so movement might be rightly restricted for reasons of safety, security, or privacy). It also means that we should assist those whose movement, health or senses are in some way hindered to help them achieve these purposes as much as possible. □

CONJECTURE 14.5.27. *A rational agent is characterised by intellect, the ability to make moral choices, and free will. The natural purposes of a rational agent are the acquisition of knowledge (to satisfy the intellect) and the ability to do what is desired (to satisfy the will).*

Discussion:
The goal of the acquisition of knowledge means more than that we have a moral responsibility to educate people, and to seek education ourselves. It also allows us to determine the manner and purpose of education. Firstly, we should distinguish between truth and falsity. To educate someone in false knowledge is not to educate them at all, for falsity is not really knowledge. It is not knowing that satisfies the intellect, but knowing the truth. This means that education should be accurate; it does not necessarily imply that it must be precise. Greater precision is, of course better, but it not always possible. It is wrong to imply that the precision of the knowledge is greater than it is. For many things, then, we should teach people that the truth lies between certain bounds, and leave them with a desire to be part of the generation that reduces the imprecision. Being overly confident in the teaching is not good; as we may leave the students certain of some error.

Secondly, we may ask the purpose of education, and I would like to comment on several goals that are often mentioned. The first is to pass exams, and gain qualifications. This is, I think, the worst reason for going through schooling. For it is to make the exam, not knowledge and understanding itself, the goal and our intellect the means. Rather, the intellect should be the goal, and the examination one of the means. If we make the exam the goal, then our desire for knowledge will be limited to what is needed to pass the exam. Yet the intellect's natural desire is unrestricted. Examinations are a tool; an important part of the processes of education, but they are not what education is for. A certificate is all very well, but means nothing if the person who possessed it forgets the skills or information that allowed him to earn it; it is the skills which ought to be represented by the certificate which will are needed in life. Rather than teaching to the exam, the goal should be to teach beyond the exam. Understanding is far better than memorisation (although some memorisation is necessary to understand). The person who merely memorises facts to pass the exam is inflexible and useless; the person who is able to understand enough to pass the examination by deducing the facts from more limited information is able to reach far higher levels. A person who dropped out of school but has understanding is better than the man with the highest degrees who unable to use them.

Secondly, education may be used to provide specific skills for a job. This is, of course, important, and given that we need to work it is important that we are able to do so efficiently and well. However, it again confuses the means and ends. It makes the job the end (an error in itself, for each type of work is merely a means to provide some good or service to the wider society), but in practice the acquisition of knowledge ought to be our goal. Certainly we should learn what we need to for our work, but our learning should not be restricted to this; it should merely be seen as one part of the knowledge we wish to acquire. I do not, of course mean, that we should tire ourselves out through study so we are unfit for work; acquisition of our knowledge is one of our natural purposes but not our only one, and its significance should not be taken out of proportion: it is a matter of finding the right balance. But we should continue to seek after knowledge, as much as is possible given the

requirements of our day jobs and our need to relax. Better still if that knowledge is related to our line of work, through not necessary for it, as it may achieve the twin goods of making us better at what we do for a living and satisfying our intellect.

No, education should be for the sake of education in itself. It should go beyond our formal schooling, and requirements for our work. We should never be satisfied with what we know; there is always more to learn.

Education, of course, is not only to gain knowledge and understanding, but also to gain skills. Skills are, of course, a means to an end rather than the end in itself, since they allow us to achieve some good which would otherwise be beyond our reach. The top priority should be skills which are necessary for education, for example: communication (including literacy), critical thinking (the ability to weigh information), logic (including mathematics and rhetoric), the ability to remember, and most importantly a thirst for knowledge and a love of the truth.

Freedom can be defined as the ability to do whatever is desired. From the natural tendency towards freedom, we obviously and immediately obtain the key liberties underlying Western Culture, including the freedom of speech, of association, of religious belief, which are restricted only by the more important need to protect rather than harm life. We have the responsibility to allow all speech baring any incitement to violence; to permit religious practice and belief except for those belief systems which encourage anger and violence, or venerate the unrepentantly violent, and so on. In particular, our modern society is starting to show an intolerance towards offensive or intolerant speech. This is wrong; it is usually not good to say something which may offend someone else's character or deepest beliefs, although there are important exceptions, such as attempts to correct them from wrong belief or bad character. However, it is even worse to restrict them. Firstly, because by allowing such people to speak, we reveal what they are and what we are, which is the first step towards correcting either them or us, whoever needs to be corrected. Secondly, it violates their rational nature: by restricting their freedom to speak, we diminish their status as rational agents; we are saying that we would rather they were unthinking machines forced to parrot our own prejudices. Given that everyone is wrong about something, if only permit speech that we like, we force people to tell untruths.

Equally important is freedom of thought. We cannot and must not tell people what to think; but instead present all views with the best supporting evidence available. Historically the religiously-minded failed in this area; today it is more the atheists and children of the enlightenment who campaign to restrict the Christian and other religious perspective in schools and force people to think like them.

Now this definition of freedom has an obvious problem: all of us frequently desire evil. To say naively that freedom of will is a good leads to many contradictions and frustrations of this good. For example, if there are two men travelling, and one wants to turn left and the other right and they desire to travel together, then at least one of them will be frustrated. More pertinently, perhaps, is the case when both the robber and the rightful owner desire to possess a particular television set – they cannot both get what they desire. It is clear that in any society where the people have contradictory desires, complete freedom is impossible.

If freedom is the power to do whatever we desire, then there are two ways in which we can solve this problem. The first is to restrict the power; we restrict the power of the burglar to rob, for example, by locking our doors. The second solution focuses on the 'desire' part of the definition. We have the thief change his desires so that he no longer seeks the illicit objects. Now, it should be clear that the second of these alternatives is far superior, because the first approach restricts a good while the second brings the various individual goods into harmony, so they can coexist together. True freedom, then, is not so much a question of what we can do but of what we desire, and in particular whether our character is virtuous.

One of the marks of a good society is that the desires of one person do not conflict with the desires of another; while this consistency is not sufficient to determine a moral

code; it is a necessary part of any objective and reasonable morality. Thus a good society is also one that is free; and the most important way to achieve this freedom is to encourage people to a virtuous character; without instruction in virtue as a primary goal in education there can be no liberty and there will only ever by anarchy, confusion, conflict and misery. The virtuous can be at liberty; and though I have not proved it I would go further and say that only a society where all are virtuous can be at liberty. Therefore the promotion of virtue is the most important step towards the goal of liberty, which is the fulfilment of the natural purposes related to the will. Of course, this is not everything; we must also seek to increase our capability in those ways which are commanded by ethical principles (such as education), or are morally neutral, or by removing restrictions (such as unjust laws or bureaucratic regulations) which hinder virtue or even promote vice.

Our world has three chief underlying problems, which affect every nation: ignorance of knowledge (or, even worse, and even more widespread, belief in things which are false), ignorance of moral understanding (or, worse, incorrect moral opinions), and defect in character. The fruits of these underlying causes are obvious and all around us: corruption, anger, detachment, violence, overreaching bureaucracy, inefficiency, laziness, imprudence, sexual disorder and family breakdown and so on. Indeed, they are not just fruits, but encourage the initial ignorance. Once we habitually start to perform some vice, we naturally try to justify it; leading us to making all sorts of mental contortions to claim that it is in reality a virtue, or at the very least not as bad as true reason would have us believe; those contortions lead us to question other parts of ethical thought, and in this way our understanding can be affected and distorted by a practised vice, leading to a vicious circle. The defects and disorders are most obvious in those in authority: politicians, business and media leaders, and so on, because their particular failings adversely affect many people and not only themselves. But it affects all of us, from every class. The lack of good character is destroying our civilisation. Knowledge, and the power it provides, without virtue is the biggest threat to our world. This, of course, brings us back to education. If the first goal of education is to teach knowledge and skills, it is perhaps even more important to instil good character and virtue, which are necessary for a functioning society. This should be done both by schools, parents and churches. Of course, our nature being what it is, such education will not be wholly successful, and everyone will still fall short; but they will fall short by less than if they lacked the education. What moral education we have now is not working; partly because it is focused on flawed utilitarian principles rather than on sound logic, and partly because we frequently encourage vice rather than virtue.

I strongly believe that proper moral education will help, but it is not the full answer. It will not lift us to the top, but it will help prevent us from falling to the bottom. The reasons for this are straightforward. Firstly, human nature is weak and easily corrupted. We are often tempted by a lesser good, which draws us away from the complete good; we are often ignorant of the harm we do to those outside our usual circles; we tend to surround ourselves with like-minded people, which serves to compound our errors; we make frequent errors of judgement. But most of all, education cannot cure all our ills because our teachers are also flawed, perhaps most especially if we try to teach ourselves. Hume, Kant, Mill, even Aristotle, all made mistakes; and I myself, of course, fall far short even of their standards. That does not mean that we cannot know some moral truth, perhaps even most moral truth, but we cannot know for certain all moral truth by our own efforts. And, of course, even if we manage to establish a near-perfect society, history teaches us that it will not stay there: moral regression occurs as frequently as moral progression. For example, slavery was rampant in classical society, outlawed by the medieval Church, reintroduced in its worst form in the early modern era, banned again but this time globally under the power of Western military might, and is now (worldwide) on the rise once more. The imperfect cannot continually make perfection; but perfection is what we need to have true freedom. There is no perfect world that can be built by human hands; history tells us that attempts

to do so invariably lead to the worst societies of all. □

CONJECTURE 14.5.28. *A society is a group of social rational agents working together for their mutual common good. Responsibility for providing each individual good is distributed between the various members of the society so it can be produced more efficiently. This requires mutual co-operation, delegation, which in turn requires communication. Given that certain tools are required for each individuals work and care of themselves, these individuals are entitled to possess these objects so that they are readily available to be used according to their proper purpose. The natural purpose of a society is therefore to provide the needs and beneficial wants of its members; the means to this end include communication; gratitude and mutual respect; justice; service; respect for property; compassion; and order.*

Discussion:
As stated above, man is also a social animal. Indeed, we must be so because of our nature. It takes a man and a woman to procreate, and a newborn child is helpless; the child needs the attention of both its parents to grow well, not to mention any teachers external to the family. Therefore the basic unit of human society is the family: husband, wife, and children; growing into the extended family when other relatives are included. The family is crucial in the propagation and continuance of the human race; not only because of its role in procreation, but because it is the best environment to raise children and teach them the basic standards needed for human co-operation and fulfilment. The husband and wife each provide a different role model to the children, and both parents are essential to provide role models of both sexes to the child, so that they can understand both what it means to be a man and what it means to be a woman. Such understanding is essential if the child is to build a successful adult life. Therefore a man cannot make a good mother nor a woman a good father.

Each family is part of a wider community. This, again, is required by human biology, since the children will need to move outside the family to find their husbands and wives. Since the community is built from families and individuals (either directly for single people or as part of a family), a prosperous society is one where its members are prosperous. The first goal of the society therefore must be to protect the family and the individuals within it, and to assist them in achieving their own individual goods.

Now, since it is difficult for each family to provide for themselves everything that they need, there is a certain diversification of skills within a community. Every type of work: every produce of industry and agriculture, requires detailed knowledge and skills, and it is not possible for a single individual to be a master of every trade. Some people provide food and water, others industrial products, others important services, others manufacture clothes, and so on. Each of these services has a benefit to the wider community; and it exists so that that benefit can be provided. We thus deduce that each company has a inherent purpose; not something defined by the people within that company nor from the attempt to make a profit, but rather an objective goal in meeting a particular societal need. Similarly, each worker within the company serves to provide a particular function needed by the company to achieve that goal. For example, a school has the inherent purpose of supplementing the parent's education of their children. Within the school, there are various roles required to achieve this goal; most obviously teachers, to provide the actual education, but also cooks to provide healthy and edible food, cleaners to ensure a healthy and safe working environment, librarians to ensure that the required educational and recreational books are available; in a modern school IT technicians to manage the computer systems; and administrators to oversee the finances, admissions and student welfare; students to absorb the lessons given by the teachers; and some others. Every member of the school's staff has a particular

purpose which contributes to the overall purpose of providing education; as do each of the students. By comparing their performance against this purpose, we can determine whether or not they are good at their job. This is the basis of occupational ethics, one of the societal goods.

One of the features of society is therefore that it is diverse, not necessarily in the modern use of the term which tends to focus on racial, cultural and sexual diversity[46], but in the sense that it is made up of individuals with different gifts, talents and skills, each of which are needed to allow the whole community function well. Each of these skills are vital to the well being of the society; the plumber and the professor are therefore of equal value. If we were all professors, then who would fix the drains? If we were all plumbers, then who would teach? The members of the society are equal in importance, but not interchangeable and not equal in every way. A diversity of skills is to be encouraged, and each member of society honoured for what they do, and trained specifically in line with their gifts.

In the modern world, of course, many of the people leading companies and organisations would replace the service to society as their primary good with other purposes, such as profit, or paying dividends to shareholders, or their own personal wages or security. Another purpose of a company could be to provide jobs and wages to its employees. There is, of course, something to be said for all of these. For example, shareholders invest in a company so that it can expand, expecting to receive a return on that investment. It would be unjust for the company not to give them a fair return; which requires that the company makes a good profit. Equally, a basic level of wages are required to provide the fundamental needs of the family. However, it is most definitely wrong for a company to focus on its own profit or excessive wages at the expense of its role for the wider society or for its workers. Money is an organisational tool used to help order the different parts of society and ease transactions between them; it ought not be viewed as an end in itself. Indeed to do so is an evil.

The purpose of money is to ease the exchange of goods and services, so that the supplier of a service and the one who demands it have a suitable measure of value so they can agree a mutually beneficial price. Wealth is not measured by money, but is instead is a measure of the extent to which somebody's needs or wants for goods or services are satisfied. Somebody who has purchased something he wants at a good price is better off than he was before the purchase. He has less money, but more wealth. The value of a product depends on its scarcity. If there is an abundance, and it is easy to obtain, then the price someone is willing to pay for it will be relatively low. If it is scarce, and it is difficult to obtain, then you will need to pay a large price to obtain it. This is true for the valuations of individual people just as much as it is for an economy as a whole. Different people will value a particular

[46]Racial diversity is morally neutral, since the different races make no difference to the common humanity; the only proviso is that given enough time, each isolated culture is likely to become racially homogeneous as there is interbreeding. Sexual diversity, in the sense that both men and women are equally essential members of society and ought to be treated as such, is clearly a key aspect of a good society. Forms of sexual diversity expressed in terms of harmful sexual attractions (as the word is so often used these days), or which deny the distinctions between men and women or which understate the similarities, are not good as they are contrary to reason. Some forms of cultural diversity, though not every form of it, should be opposed as they may obstruct the shared purpose and confuse the shared values that allow a society to function efficiently. This depends on whether the differences in culture are accidental to humanity's good or relate to matters essential to it. For example, some cultures and religions, the extreme cases promoting violence, slavery, mutilation, sexual immorality and so on, are certainly contrary to the goods of humanity and should be vigorously opposed and those cultures suppressed if they cannot be transformed. Other cultures are more benign, though still require improvement, since no human culture is perfect, nor can any be perfect. It is important to note that human cultures do not naturally progress in morality: the natural direction of their changes, except where a great deal of effort is put in, just as often goes in the other direction, and has done throughout the history of rising and falling civilisations. Change is rarely for the better; change in the moral code of a successful society is almost invariably for the worse. However, good cultures need not be monolithic: there may well be different ways of organising a good society. If so, then these should be allowed to grow side by side - though not necessarily mixed together, which might lead to a compromise between the two cultures, and compromises between two goods are usually worse than either of them.

good or service differently, depending on how much they want it, and how easy it is for them personally to obtain. The consumer, who has a surplus of money and a deficit in the good or service he wants to buy, regards the product as being more valuable to him than the money (up to a certain limit). He values the product more than he does the price he pays for it. The producer, who has a deficit of money and a surplus of the product – more than he can personally use, is in the opposite position. He values the money more than he does the product. Thus because the need for the product differs between the consumer and the producer, they have different internal valuations of its worth. In this way, a fair and mutually beneficial price can be agreed in which they both walk away from the transaction having gained. But there is only a narrow window of mutually beneficial prices. Only the consumer and producer know how much the product is worth to them. Anyone else is just making guesses, and will certainly guess wrong at least some of the time. If there is a trend of pricing the product too high, yet forcing it to be bought, then the consumer slips into poverty. Price too low, and the producer's business is no longer sustainable. But money is worth less to the person who has a surplus of it, while it is valued more by the person who has too little. Thus in an unfair transaction, somebody will gain a certain amount of money above the fair price, and the other person will lose the same numerical amount. But, if there is a consistent trend in who wins and who loses, the one who gains needs the money less, and thus values it less, than the one who loses. In this way, overall the wealth of society decreases. Because of the way the economy is interconnected, if one person fails, then others will suffer as well. A centrally controlled economy will thus always suffer economic collapse, because economic knowledge is necessarily dispersed across the entire population, and no small group of self-proclaimed experts can ever know enough to consistently make the correct valuations. There will be a trend of unjust prices, and then some people will suffer, and then all people will suffer.

The goal of any economic life should be the creation of wealth. Wealth here means those things which people want or need to fulfil their natural ends, such as food, clothes, refrigerators, tools and so on. Wealth is thus created through the production of goods and services, in response to demand. Its value arises from that demand and varies according to the demand. Wealth is consumed when those goods and services are used up. The goal of any individual, family, community or nation should thus be to produce at least as much wealth as is consumed. Poverty arises when consumption of wealth exceeds its production. Money is not wealth, since it does not directly satisfy any natural end, but only indirectly as a means of exchange. However, when people treat the acquisition of money as the end, then people who are capable of it tend to seize more for themselves than is deserved in proportion to how much they provide to society. Thus the purpose of money is seen not as providing valuation but instead something to be obtained by me. In this way, what people desire from money differs from the purpose that defines money, and this is an obvious evil.

Overvaluing your own worth to society, which implies taking more from an exchange which is just, necessarily means undervaluing someone else, so that the other party comes off worse in the exchange. In principle, this should not happen. Competition means if you do not offer a fair deal, your customer will go to someone who will. Prices are negotiated so that both parties gain from the transaction. In practice, it doesn't always work out like this. Temporary imbalances in the market mean that some people have more power in the exchange, and are able to dictate terms. There can be cartels, where all the producers agree to charge a high price. A salary is also a market exchange; the employee is freely exchanging something they have (namely their labour) and which the employer wants for something the employer has (namely a livelihood) which the employee wants. In principle, this exchange comes from a free decision by both parties, and thus of mutual benefit. In practice, not always. A cartel of workers (a union) can lead to the labour being over-valued, to the detriment of the company and ultimately the customer and thus the society as a whole. On the other hand, an oversupply of unskilled labour will leave the employer with

the stronger position, leading to salaries suppressed to the point where they can no longer cover the cost of living.

Abolish the markets? But without the market, there is no just way to know how to provide a valuation, since value depends on the meeting of the demand of individual people with the suppliers, and the market is just this meeting of demand and supply. Any government or other agency which establishes price controls will get it wrong. The economy is just too complex, too fast changing, and requires too much specialised knowledge across a vast range of fields for any one person or group of people to cope with. The knowledge is distributed across the whole population; therefore economic decision making also needs to be distributed to individuals. Centrally controlled economies are inefficient and cannot correctly determine value. If they act on behalf of the workers, then the workers will (being selfish) demand more and produce less. This invariably leads to greater consumption than production, and the society itself will collapse into poverty. Instead of having some people poor (as happens where there is a free market), you end up with almost everybody poor.

Keep the free market? Better than trying to control it, as history clearly shows. But still far from perfect. Society as a whole will be richer, but the over-supply of low skilled workers means there will always be the poor and an unjust distribution of wealth. I do not think that there is any economic structure which guarantees that nobody is left in poverty. It is an insoluble economic problem.

But it is an economic problem that is caused by a moral problem. The problem is that people over-value their own worth and under-value others. This is greed. There is enough to satisfy everybody's needs. The problem comes when people seek their own money as more important to them than providing for the worse off in society, either through philanthropy, or simply being willing to pay a little more – a pittance to them, but the difference between life and death to those less well off – in economic transactions.

I do not wish to discuss economic systems in detail here. No system is perfect; indeed it seems to me that we have to decide which is the least bad of several very lousy options. I have my own views about what that is, but as I have not studied economic theory my view does not count for much, so I will not share it in detail. My concern is for the ethics of the situation. It should be seen that the primary moral purposes of the company are threefold: firstly to employees – that they might fulfil their obligation to labour for society and also receive the means to benefit from the labour of the rest of society; secondly by providing a service and benefit to the rest of society, and in particular their customers; and thirdly by providing the means through which the work of their suppliers can fulfil its purpose of benefiting the wider society. The system breaks down when the leaders of the company substitute other goods for these primary goods. This is not so much a failure of effort; but of having morally wrong priorities. Economic failure is therefore first and foremost a moral failure: a lack of virtue at all levels of society, characterised by people failing to work for the purpose that they ought to be working for. No economic system, no matter how well designed, will prosper if the people in positions of power are corrupt. No economic system, no matter how poorly designed, will fail if the people throughout the society are good and work hard for the purposes they ought to be following. The problem facing economists is to study how to distribute finite resources given infinite human potential for corruption.

Mankind's primary purpose has nothing whatsoever to do with money. Earning money is not in itself a natural purpose of man. A man's worth is not judged by the size of his wages or the quantity of his possessions – quite the contrary. Men who make do with little are often vastly better as people than those who lavish in plenitude. A certain amount of money is necessary in our world for survival; a little more than that to live a comfortable life. But these are simply instrumental goals; to provide for other needs. That does not mean that earning in excess of this is in itself an evil; although it may be an evil if it is

hoarded when it could be used to support others in desperate need[47]. But setting money as a personal goal, as though it were a natural purpose of man – one could call this the love of money[48] – is a great evil, and the cause of much misery in the world.

This does not mean that there should be no increment in wages; everyone should receive the same salary. On the contrary, justice demands that those who do the best job according to their role, and who work the hardest should be rewarded the most. Equally, those in the more skilled and harder jobs might be expected to earn a little more than unskilled labour, as should those who have given longer service. This does not, of course, mean that salary should be based on the rank within a company. But justice is focused on everyone receiving what they deserve. It is unjust to pay a worker less than their work is worth – whether that worker is at the bottom or top of the company – and it is unjust to pay them more – again, whether they are in a high or low position. Of course, worth is a difficult thing to measure, although it must include how well the job is done. The rank within the company is of considerably less significance. Baring trainees, all employees are equally necessary for the company's success (otherwise they would not need to be employed), and all should be justly rewarded.

The next necessary societal power I will highlight is communication. I think that most people will agree that communication is essential for a well ordered and complex society; or at least one where its members are required to go beyond mere instinct. The purpose of communication is to convey knowledge to another. This may be to teach; or to outline a future plan; to entertain; to warn or rebuke; to come to a compromise or agreement; to barter; to coordinate efforts in a single task; or a host of other reasons. For most of these reasons (entertainment being the only exception I listed), we see that accuracy is essential for the communication to meet its purpose to build up society successfully. When people lie, they do so for their own benefit (or what they perceive as their own benefit) and at the expense of others. The purpose of communication is to convey knowledge, and falsity is not knowledge (one exception is when you are telling a story for someone's amusement; but here everyone knows it is a story - at least by the end of it - and therefore no-one is deceived and there is no intention to deceive. Indeed, the best stories convey some more fundamental message about human nature which can be used to teach). Therefore a good member of society is one who by nature speaks honestly; which in turn implies that he thinks honestly. To lie is to desire that the other person remains ignorant of the truth; to know less than he can (worse: to understand falsely); to only have a limited mind incapable of being aware of some particular knowledge; and therefore to desire that the other person is in one particular way less than human, since to be a rational animal is to have the natural purpose of understanding every truth it can understand. □

DEFINITION 14.5.29. **Agape**:
Agape (often translated as love, although this is misleading because the English word 'love' covers many different senses, some of which contradict agape in certain circumstances) is the desire for goodness in itself. We have already seen that people ought to be good; agape is the alignment of the will with this reality. It means that we seek to strive not only for our own good, but also the good of others. This is the highest and most important virtue, obviously, since unless we desire goodness, we will not desire any of the other virtues (or seek them not because they make us good, but for some other purpose which will ultimately

[47]Although I am against simply giving away cash to those who are capable of earning it – that only encourages laziness and a dependency mindset. In the long term, it degrades a man rather than building him up. I am discussing those who due to misfortune, disability or natural disaster are unable to provide for themselves; those we all have a responsibility to help.

[48]Which is one form of the vice of greed.

corrupt them); and if we do not desire them, we are most unlikely to acquire them, and if we do, by chance, get hold of one we would be just as likely to cast it aside. △

DEFINITION 14.5.30. **Temperance:**
Temperance is the desire to take substances, such as food or drink, in the right quantity. As a living organism, a natural purpose of the body is to be able to sustain itself, and we call this state 'health'. Both to over-eat and to under-eat are damaging to our health, and thus go against our natural purpose. In the middle is the optimum point which comes closest to fulfilling the natural purpose of maintaining health (although, of course, our mortality means that no diet will perfectly maintain health, but there is still a maximum), and this defines temperance. △

DEFINITION 14.5.31. **Fortitude:**
Fortitude is moral courage; the strength of character to stand firm for what is good when under pressure, either from ourselves or others. Clearly a lack of such character is likely to encourage us to slide into evil one way or the other, and thus fortitude is a virtue. It lies between the vices of cowardice and recklessness. △

DEFINITION 14.5.32. **Prudence:**
Prudence is accurate knowledge of moral principles. This is a virtue of the intellect. The natural purpose of the intellect is to seek right knowledge, and prudence is right knowledge and its lack either ignorance or wrong knowledge. Thus prudence is a virtue. △

DEFINITION 14.5.33. **Justice:**
Justice is the desire to treat people as they deserve. It lies between the extremes of treating them too harshly and too leniently. It does not mean that all people should be treated alike. It ultimately arises from agape, indeed it is one of the practical applications of agape. For agape means to desire goodness in general, and part of this involves a desire that specific instances of goodness should prosper. In terms of people, it means that it should be desired that people who exhibit goodness should prosper materially and in other ways, while, in comparison, people lacking goodness should not prosper so much. This is one way of glorifying goodness. Obviously, there is a desire that all should become good; but the reality is that not everyone is good (indeed, nobody is good - including myself - there are only degrees of evil). Justice should, of course, be tempered by the notion that we should help the vicious become virtuous, and rewarding the good and punishing the evil might seem to be a way to encourage this. But this is wrong; true virtue is not achieved by hope for reward, nor loss of virtue avoided by fear of punishment. Virtue is only achieved when we make the desire for goodness an end in itself. Justice is not a cause of people's improvement, but rather an effect of their failure, or, indeed, success.

The natural tendency of a man (or woman) in a state of *eudaimonia* is a lasting contentment, as stated earlier. Not the fleeting pleasures of vice, which pass away almost as soon as they come and require ever higher doses to get the same 'reward', until they become a slave-master and despair for the vicious. Instead, the virtuous man lives according to his natural purposes; his desires and being are in full accord, and thus there is no internal

tension within him. His desires and purposes (which are focused on the goodness of his character) are continually fulfilled, and the natural consequence of this is a lasting joy. This joy can, of course, be frustrated by external circumstances, like all natural purposes; but like all purposes it is good for it to be fulfilled. Thus it is good for a virtuous man to be comfortably and securely content. It is good to reward the virtuous; to help them be content.

But what about the vicious man (or woman)[49]? Here, quite clearly, the opposite is true. If the natural consequence of *eudaimonia* is happiness (though virtue, rather than happiness should be our goal: to seek happiness in itself is an evil because it means that the focus of desire is not pointed towards *eudaimonia* but something else), then the natural consequence of vice is its opposite. His desires and being are in discord; the natural end of that is a life of misery. As with all ends, goodness is achieved when it is fulfilled. Thus it is good to punish the wicked, removing the barriers to the fulfilment of the natural tendency of his state. That is not to say that we are to induce suffering for the sake of suffering; but we are to seek that a deficient character does not evade its natural consequences. Obviously it would be better if the man become virtuous, so his beings and desires are in agreement, and his natural end is no longer a life of suffering but of joy. But if that does not happen, then punishment of the wicked, in due proportion to their wickedness, is not an evil but a good because it is fulfilling the natural purposes of a vicious character.

This, then, is the virtue of justice: to seek that the virtuous are rewarded in proportion to their goodness, and that the wicked are punished in proportion to their vice. It is wrong to reward the wicked, or even the partly good more than they deserve; it is is a vice to punish the wicked both too little and too much. But there is, somewhere in the middle, the right balance: finding that balance is, of course, a lot harder to put into practice than acknowledging that it exists. Δ

DEFINITION 14.5.34. **Hope**:
Hope is defined as the reasonable belief that the future (or one aspect of the future) will be better than the present, or remain as good as the present. It means to look forward to a good future. This, unfortunately, is one word where the English meaning has changed over time, so the modern understanding of the meaning is not the same as the older traditions which I am drawing from. The modern understanding of 'hope' is one where that future is particularly uncertain – we use the word to imply uncertainty. However, I will here use the word in a more traditional sense (more in line with the Christian understanding of the term) where the word contains implications of assurance and confidence.

Hope begins with a promise that there is a good which may subsequently be actualised and ends with either the fulfilment of that promise, or the situation where it is impossible to fulfil. It is impossible to hope for something we already have. It is an intellectual virtue because it means belief in the fulfilment of promise of someone reliable and benevolent. The negation of hope is disbelief in something reliable, which is clearly an intellectual error.

Hope lies between two vices; the first is despair, the unreasonable belief that the future will be worse than the present (or the same as a bad present), and the second is unrestrained optimism, the unreasonable belief that the future will be better than the present. In the first case, the despair leads us to not work towards the good, meaning that we are either unprepared for its realisation, or that, of it relies on our efforts, it will not be realised at all. Secondly, unrestrained optimism leads us to waste our energies working for or preparing ourselves for some impossible good, when we would be better served by trying to work for some good which is achievable. Of course, unless the hope is certain, when we hope we

[49] I use vicious in its original sense of 'deficient in the virtues' rather than the modern sense which focuses on the vice of cruelty.

should also be prepared for the eventuality that that hope will fail. △

DEFINITION 14.5.35. **Faith:**
Faith is defined (in the first sense) as a well-informed belief and trust in a theory or person, and a willingness to act based on that trust despite risks if the trust should be misplaced[50]. 'Well-informed' in this context means informed primarily through accurate observation and experience, either our own or from other people we have good reason to trust.

In particular, faith looks to the future, to what has not yet been experienced. It is a trust that a person will act consistently with past promises; or that natural law is ultimately consistent so theories built on prior evidence will prove true when put to practical test in circumstances not yet experienced. For example, Christian faith is to trust that God will fulfil the promises He made through Jesus Christ; an engineer's faith is that his bridge will not collapse when it has to bear its maximum designed load. Both are built on evidence; and both are held with surety on the basis of that evidence; but neither are yet directly tested (though they one day will be).

Faith, as the Christian defines and uses the word, is not contrary to evidence but rather a response to evidence; and from that foundation extends the knowledge gained to new circumstances where it is willing to step out trusting that the universe is consistent and rational. Thus faith, understood in this way, is at the heart of the scientific method.

Genuine faith is an intellectual virtue; it involves putting your trust in what is true. To not put your trust in the truth is a distortion of the purpose of the intellect to seek truth. △

My intention has been to outline the basis of natural law ethics rather than derive the virtues and vices, so I will end my discussion here. Those interested can consult the relevant literature (Fagothey, 1959; Rice, 1999; Devine, 2000; Gomez-Lobo, 1999; Oderberg, 2000). See (MacIntyre, 2007, 2006, 2013, 2016) for a more general defence of virtue ethics.

14.6. The Goodness of God and Man

CONJECTURE 14.6.1. ***God is necessarily good***

Proof:

(1) Anything that is essentially good is good necessarily (i.e. its goodness cannot be created or destroyed, unless the substance is created or destroyed).
(2) God is essentially good.
(3) God exists essentially.
(4) Therefore God is necessarily good.

■

Discussion:
The second and third premises were shown above (conjecture 14.5.16 and definition 13.1.12). The first premise is clear: if something is part of the essence of a being, then if the being exists then so must that predicate. Since God is essentially good, and exists essentially so

[50]A second sense of faith, not discussed here, is of moral constancy; for example maintaining what we know through reasoned and reliable observation in the face of emotional pressure. A third sense, again not discussed here, is of a general system of belief, in particular religious belief.

God Himself cannot be destroyed, it follows that God is necessarily good. □

Objection:

(1) To be perfectly good means to fully possess all the virtues.
(2) Courage is a virtue.
(3) God is perfectly good.
(4) God must therefore fully possess the virtue of courage.
(5) Courage, the overcoming of fear, is only possible if there is a genuine risk of harm or danger.
(6) There is no risk of harm or danger to God.
(7) Therefore God cannot possess the virtue of courage.
(8) The contradiction between points 4 and 7 implies that one of the premises is false, most likely that either God is not perfectly good, or indeed that God does not exist.

•

Response:
Goodness for each type of being involves being fit to fulfil its natural purposes. Given that the natural purposes of each type of being are different, goodness for each type of being might be manifested in different ways. This means that the virtues for one type of being might differ from the virtues for another type of being. God, like us, possesses an intellect and will. Therefore those of our virtues which arise from the fulfilment of the natural purposes of the intellect and will (such as justice and prudence) will be present in God; not precisely manifested in the same way as our virtues, since God's intellect and will are not identical to our own, but certainly analogous to them. However, other of our own virtues are derived from the natural purposes of our bodies (such as temperance and courage), and these have no analogue in God, since God does not possess a body. Therefore the first premise above should read 'To be perfectly good means to fully possess all the virtues applicable for the being involved.' The second premise would then be for a valid chain of reasoning 'Courage is a virtue applicable for God.' Yet this statement is false, as we have seen. Thus there is no reason based on arguments of this type to question the premise that God is good, or that God exists. ○

DEFINITION 14.6.2. **Hate:**
To hate something is to desire harm for it; that is that it should move away from its state of goodness. This may include either that we desire that the object of our hate ceases to exist, or that its nature is corrupted so that it is less able to fulfil its natural ends, or that it is maintained in a state of corruption.

To hate evil means to desire that that evil ceases to exist. △

CONJECTURE 14.6.3. ***It is impossible to practice agape without hating evil***

Proof:

(1) To agape something is to desire that it becomes good.

(2) For something to become good requires that the both the uncorrupted parts of its nature are enhanced, and that the corrupted parts of its nature (those elements which are not directed towards its natural ends) are destroyed.
(3) To agape something is to desire that the corruption or evil within it is destroyed.
(4) To agape is to hate evil.

■

CONJECTURE 14.6.4. *God necessarily fully embodies agape*

Proof:

(1) In designing creation, God starts by establishing the natural purposes, and then constructs the forms.
(2) Everything is created in order that it might fulfil its natural ends.
(3) God's desire for a being is that it should fulfil its natural ends.
(4) Therefore God's desires that every being becomes good.
(5) Therefore God's will is necessarily fully in accordance with agape.

■

CONJECTURE 14.6.5. *No living human being is good.*

Proof:

(1) To be good means to be fit to fulfil one's natural purposes.
(2) One who is good is incapable of acting in a way contrary to the moral duties derived from those purposes.
(3) Observation tells us that we have all fallen short in one way or another,
(4) Therefore every living human being is evil.

■

Discussion:
I am, of course, using the word 'evil' in its technical sense of deviating from goodness to any degree. This use is reasonable, because the goal for God is to create a perfect society according to His purposes. This society will, of course, be eternal; which means that even the smallest flaw will eventually emerge and cause imperfection.

As for us all falling short in some respect, that is obvious. Ever lied? Ever got angry? Ever felt a twinge of lust? Ever stepped on an ant in a vindictive manner? Even if you have not, even the capacity to do so is enough. Is this standard harsh or unjust? No; it is logically necessary. The standard we ought to aspiring to is simply too much for anyone to achieve; at least not without assistance from God which requires firstly our acceptance of our failings. Many do not feel the force of this idea, because to experience it we must be lifted enough away from sin to recognize the tension. To cite an analogy from C.S. Lewis, a man underwater cannot feel the rain. Or in the language of Paul, many people are so conformed to the desires of the flesh that they do not recognize the flesh as being separate from themselves; or they believe that the desires of the flesh are good. Of course Christians are not immune from this flaw, and there are doubtless aspects of my life where it applies as well (one can recognize flaws in one aspect of our character, while being oblivious to

them in another), although every Christian is aware of the tension between the desires of our rational nature and our instinct for immorality. But many people, both Christians and others, who have tried to live a life of virtue will recognize the force of this tension. We know something is wrong; but we cannot control ourselves: we still do it, as though it were an addiction.

Of course, this doctrine, in focusing on our capacity for evil, and not just deeds we have committed, we do not restrict the condemnation to just those who have reached adulthood. This idea relates to the fact that we are conceived with the capacity for error; a capacity that then flowers as we are raised by our imperfect parents, and enhanced even more when we go out from our family and into the wider world. This imperfection is not only in our character, but in the imperfections of our bodies; our genetics allowing us to grow in this way. We inherit it from our parents, both literally and through imitating the mistakes they make as they raise us; we acquire it from imitating the goals and aspirations of the wider society; often we do not realise how deep the corruption is within us. While this corruption, this capacity to be less than perfect, remains, we are not fit for heaven. We are therefore destined for hell from our conception.

Everyone therefore falls short; we all have the capacity to miss the mark of virtue and we all do miss the mark. We all fall short of God's perfect purpose for us. I do not exclude myself, or any Christian from this. If I label someone as a sinner, I do not mean that they are worse than anyone else, or that the error I am pointing out is worse than the flaws within me. They may well be better than me, and often are. But I am saying that they fall short of the standard, as do I. It always amazes me that the Christian doctrine of original sin is so despised by the world, when the evidence for it completely surrounds us, and we see it clearly within ourselves. We do not teach babies to lie, or scream for themselves, yet they do it. We do not teach politicians to be corrupt, or CEOs of multinationals or leading sportsmen and actors to pay themselves far more than is just, but they still do it. We do not teach the members of societal movements to rage against injustice while being unjust themselves; yet they are. This capacity is part of us all; therefore it is something we inherit either as part of our alongside our human nature. Of all the key doctrines of Christianity, original sin is the most easily confirmed; yet people do not see it, perhaps because they do not understand how high the standard is, or perhaps because they do not want to see themselves as condemned and evil. □

CONJECTURE 14.6.6. *Being good by society's standards is not good enough for God.*

Proof:

(1) God's aim for humanity is a society characterised by an eternal harmony between people, reflecting the harmony of God.
(2) Full harmony between people is only possible if the people are perfectly virtuous.
(3) Full eternal harmony is only possible if people are incapable of vice.
(4) Society's requirements fall far short of perfect virtue; excluding only major vices rather than the less severe examples of vice. Indeed, some states of virtue are frequently dismissed as being naive, impossible, or even treated as evils by elements of society.
(5) Since society's standard of morality is less demanding than God's, it is possible to be regarded as good by society while still falling short of God's standard.

■

Discussion:
The premises of this argument follow from the discussion above – God's purposes for mankind are identical to our natural purposes, and we fall short of these (conjecture 14.6.5). We can observe what standard of ethics is demanded by wider society. Nobody whom we label as good is perfect. Indeed, we are unwilling to set moral perfection as the legal standard, because we are not able to achieve it and most of us do not desire to achieve it (including those who make the laws or write textbooks on ethics). □

CONJECTURE 14.6.7. *Our 'natural' desires are no judge of goodness.*

In this conjecture, *natural* should be taken to mean those unconscious desires which drive us in, for example, our emotional search for pleasure and happiness. They should be contrasted with those consciously chosen desires which follow from bending the intellect to follow natural (in the sense of following rationally from the basic facts of humanity's formal, material, efficient and final causes, and the workings of the rest of the universe) law and through training. A unhealthy 'natural' desire might be the glutton's lust for food, or the adulterer or unmarried man's lust for sex; which both come to them unconsciously and usually without any conscious choice but are opposed to the principles of natural law.

Proof:

(1) The actualised potentia of each human being differs from those potentia able to fulfil the natural purposes.
(2) The state of each human being not only differs from this ideal, but is not protected from moving away from it at any time.
(3) Any aspect of an individual human's nature can fall short of the ideal.
(4) One's character is an aspect of an individual human nature.
(5) One's unconscious desires are determined by the character.
(6) One's unconscious desires can fall short of the ideal.
(7) Unconscious desires do sometimes (though not always) fall short of the ideal.
(8) Unconscious desires may be orientated towards either good or evil.
(9) It is impossible to judge whether such a desire is for good or evil by reference to the desire alone.
(10) Natural desires are no judge of goodness.

■

Discussion:
This is arguing against the position of voluntarism, which prioritises the will over our reason, and, in the case of ethics, states that our own individual will is the guide to what is and is not good. This position is clearly flawed, because the primary question is whether or not our will is good, and obviously our own will cannot be used as the measure for that without descending into circularity. The only way voluntarism could work would be if the will was infallible. Since we know that it is not, since there is such a thing as human evil, what we happen to desire cannot be used as a standard of goodness.

The first few premises follow from the discussion above, and that our desires may be corrupt as well as the rest of us is clear. The argument for moving from saying that desires can fall short of the ideal to do in some cases fall short of ideal is that, given a large enough sample, anything that can happen will happen; and the assumption (borne out

by observation that there are gluttons, adulterers, etc. in the world) that the chance of deviation is sufficiently high that given the population of mankind throughout history it would have happened. Since there is no difference in origin between a good desire or bad desire, whether a desire in itself is good or bad cannot be judged except through some external measure, such as natural law theory. For example, one man might desire to go left and another otherwise identical man to go right, or vice-versa. Given this information, there is no possible way to judge whether one is good or the other or both or neither. However, once we know the purpose of the men's journey, to go to a certain place, then we can say (for example) that to go left is better than to go right. One woman may naturally desire to eat and another to not eat. Without the context (i.e., what the food is, and how well nourished they are) we cannot say that these desires are good or evil; but once we learn that the first woman is a glutton and the second anorexic, both are evil. We cannot judge whether our desires are aligned with our natural purpose given our form on the basis of the desire alone; we must also know the natural purpose and, in some cases, the actual form of the being. From this, the conclusion follows.

If 90% of people act in a certain way, that does not mean that that action results from a good character. Our deepest longings, temptations, passions and orientations can be and frequently are evil. Morality is not decided by majority, or even minority, vote. □

CONJECTURE 14.6.8. *No member of our race, from the newly conceived to those on the point of death, is good enough to meet God's standard by themselves.*

Proof:

(1) God's standard requires us to be incapable of committing evil.
(2) All of mankind has the capacity to actualise some potentia which is away from the ideal target.
(3) None of mankind, even if they have in practice done no wrong, is incapable of doing no wrong.
(4) None of mankind meets God's moral requirements.

■

Discussion:
The first premise was shown in the discussion following conjecture 14.6.5, while the third follows from the second, while the fourth from the the third and first. The second premise was shown in conjecture 14.6.5. The form of a good being is a subset of the form of a being; and the being may actually become described by any of the Hamiltonians described by its form. Therefore it is capable of being described by any of these Hamiltonians. Furthermore, the form of a living being contains the continuous set of Hamiltonians that give a viable creature. For a being to be naturally incapable with evil, the form of the good being must be identical to the entire form of the being; but this is impossible (certainly for us), since it is always possible to continuously distort from a state of goodness to a state of evil. The only way for a being to be eternally incapable of evil would be if some external power assisted them (or compelled them) to always remain good, as a pot holds water in a certain locality and prevents it from flowing out to form a puddle (which it would otherwise do), as discussed in conjecture 14.5.15.

This conclusion has consequences which naturally revolt us. We view a newborn baby as innocent, because they have not had the opportunity to do wrong. This argument suggests that they are, in fact, doomed to hell because they nonetheless retain the capacity for wrongdoing. We must, however, ask whether whether it is our natural revulsion which is

incorrect, or the argument which is incorrect. Unless a strong argument can be produced to the contrary (i.e. that God does not require a conscious choice on our behalf during this life to allow Him to make us good and compel us to remain there; the first premise) logic compels us to conclude that the argument is correct and our revulsion misplaced.

Now, because a newly conceived fetus is naturally doomed to hell if its life were to end at that point does not mean that God will compel them to hell in practice. There is still the idea of God's grace to consider. But we should not assume that because they seem innocent to us, that they will be regarded as so by God. □

CONJECTURE 14.6.9. *There is nothing we (as the human race) can do without God's intervention for any individual to reach a state of perfect virtue.*

Proof:

(1) To reach and remain in a state of perfect virtue, one must either naturally be forced there or be forced to remain there by some outside will (which may be one's own will).
(2) No human being is naturally incapable from straying from the state of perfect virtue.
(3) Only something perfect in itself can force something to remain perfect.
(4) No human being is perfect in him or herself.
(5) No human being can force another human being to become or remain perfect.
(6) We (as a species) cannot remain in a state of perfect virtue by our own efforts.
(7) A man-made utopia is impossible to achieve.

■

Discussion:
By being naturally forced to remain in a state of virtue means that the form of a good being must be identical to the form of the being. We have already suggested that no human being meets that target (conjecture 14.6.5). The third premise follows from conjecture 14.5.15, and the fourth premise follows from the second. That we cannot force another human being to remain perfect follows from these two premises (indeed, observation shows us that we cannot even compel someone to remain in even our own flawed standard of goodness, let alone perfect goodness), and the conclusion follows directly from the premises. Therefore, we cannot even make ourselves good, let alone anyone else.

I should state, however, that even if utopia is impossible to achieve, that does not mean that we should not strive to approach it as close as we can. Even though we will fail, we will surely come closer if we try. Of course, such an attempt would only be even partially successful if those leading it have their principles and understanding of human nature as correct as we can manage. This is why revolutions such as the American revolution (the best of them), to the French and the various Communist revolutions (the worst) have all failed; and for the most part created worse societies than those they supplanted. In the very act of using violence means to overthrow an authority, they show themselves to be flawed. □

CONJECTURE 14.6.10. *If we can only rely on our own efforts, we are all doomed.*

Discussion:
This follows from the discussion above (14.6.9), plus the definition that to be condemned to

viciousness is to be doomed. But the point is worth emphasising. In fact, it is the most important difference between Christianity and any other religious or irreligious creed. For the Christian, all of these religious efforts are futile. That is not to say that they cannot bring some benefits - there is much to admire, for example, in traditionally Buddhist societies or societies built around the teaching of Confucius, but, as any survey of the history of China, Japan and Korea will testify, even these, the best of the human attempts to teach and reach virtue, have not had any complete successes and fall well short of what is needed. Certainly, Christianity has its own rituals, and the church expects its members to behave virtuously, for example through charitable deeds. However, it is key to Christianity that none of those rituals or acts by themselves gives us any credit or enough credit[51] with God to bring us into our intended standing with him. Other religions judge people on their deeds and character; for the Christian no deeds or virtue can be good enough. That does not mean that human perfection is impossible; merely that it relies on God rather than ourselves. □

CONJECTURE 14.6.11. *If we are judged as fit for an eternal good society, it is wholly unmerited.*

Proof:

(1) An eternal good society requires that its members are incapable of vice.
(2) None of us are incapable of vice.
(3) None of us deserve on account of our character, orientations or actions to be judged fit to enter an eternal good society.
(4) If we are judged fit to be part of such a society, it is wholly undeserved.

■

Discussion:
The second premise follows from the previous discussion (conjecture 14.6.5), while the first is obvious: vice is a deficiency in goodness, and an eternally good society must therefore eternally lack vice. If the members of that society are capable of vice, then given enough time they will be vicious. For the society to be eternally good, they must therefore be incapable of vice. The third point extends this to the point of deserving; to deserve means to be entitled on account of either who we are or what we have done. But one cannot be entitled to something which is opposed to our actual (not potential) nature. One cannot be entitled to something if, once one has received it, it ceases to be what it was. For example, those people who have engaged in several years of postgraduate study, submitted a thesis and successfully defended it are entitled to receive a doctorate; because a doctorate (excluding honorary degrees) is a mark in recognition that somebody has achieved that (and is therefore capable of achieving that). If we routinely give doctorates to ignoramuses who could not even complete school, then the qualification loses its meaning. People are not entitled to a doctorate unless they have completed the work that required to earn one. Given that our actual characters and actions are contrary to the idea of an eternal good society, then as soon as we enter it, if we enter it as we are, it will immediately stop being such a society. Therefore we have no entitlement to enter such a society; and therefore we do not deserve to. If we do not deserve to enter it, then clearly if we do enter such a society it is undeserved; and if undeserved, then unmerited. □

[51] Crudely, 'any credit' represents the Protestant perspective; 'enough credit' the Roman perspective, although in practice both positions are far more nuanced, varied and detailed than I have presented here.

CONJECTURE 14.6.12. *We all naturally deserve condemnation.*

Proof:

(1) There are two states of man in the final judgement; fit for the city of God, or not.
(2) Those who are not fit for the city of God are found lacking in the judgement.
(3) To be condemned is to obtain an unfavourable judgement.
(4) None of us are naturally fit for the city of God.
(5) We all deserve condemnation.

■

Discussion:
In the previous discussion, it was argued that we are not naturally entitled to the city of God (an analogy for the perfect society which God is seeking) on account of who we are; because as soon as we enter it (unless we are first changed) it would cease to be what it is. This, however, is not true for the other destination; the worldly city. It is made of people like us and for people like us; people who deserve to wallow in misery because they naturally cannot obtain what is required for perfect happiness. By the same reasoning as above, we are 'entitled' to condemnation, and thus deserve it. □

DEFINITION 14.6.13. **Grace**:
Grace is defined as an unmerited favour. △

CONJECTURE 14.6.14. *If we are allowed to be part of an eternal good society, it is only by grace.*

Proof:

(1) We may enter an eternal good society either by merit or by grace.
(2) We cannot enter a good society by merit.
(3) We may only enter such a society by grace.

■

Discussion:
The first of these points is clearly a true dichotomy, since some outcome can either be merited or wholly or partially unmerited; and if unmerited then by definition it is due to grace. The second point follows from conjecture 14.6.12. The conclusion necessarily follows.

This is a key difference between Christianity and almost every other religion and philosophy. For example, in Islam, the good and evil (obviously defined in terms of submission to the Islamic god and his commandments) that one does is weighed against each other, and if your good outweighs your evil then you have a chance of entering paradise (although it is not guaranteed unless you undergo Islamic martyrdom – a very different concept from Christian martyrdom). In Buddhism, one reaches Nirvana through meditation and self-denial. In paganism or Hinduism by offering to the Gods. In Marxism, by abolishing class systems, private ownership of the means of production, and consequently submitting to the ruling elite of the state. In all cases, it is possible to achieve it by your own efforts. In

Christianity, no good that you do makes a single jot of difference. If you give all you have to the poor, devote your life to good service, are pleasant and nice to everyone, it counts for nothing because it is still not good enough. Indeed, because such works proceed from a corrupt motivation, they are counted as sin.

This does not mean that doing such works is wrong in itself, merely that the reason we do so should be correct, and in particular not in the hope of earning some reward, whether in this world or in some future heaven.

Of course, the question remains what the mechanism by which that grace is provided is. Or even if God has provided us with such a mechanism. That question, however, lies beyond the scope of this work. It is not something which can be deduced from reason and science alone. Our need for grace, however, is evident from reason alone. □

14.7. God and Evil

Before I move on to the actual argument from evil against God, there are a few general points to make. The first was alluded to above. One of the premises of the argument from evil is that God is perfectly benevolent. But this is not such a simple premise as it appears. How we define good and evil will affect what we mean by the word benevolent. So we can easily come up with some event which on some construction of morality is an unjustifiable evil, and will therefore seem to contradict the notion of a benevolent omnipotent God. But then the theist will respond that the benevolence associated with his God is associated with a different definition of good and evil, one in which the event in question is either not an evil or is in some way justifiable. For example, the evil might lead to a greater good later, or perhaps prevent an even greater evil. And to argue against that last point means that the atheist not only needs to be able to identify evils, but also decide on a way of ranking between them. The theist can argue that the example doesn't conflict with his God, because his God regards one type of evil far more seriously than the atheist supposed. For example, one can't just equate evil with suffering, show there is suffering in the world, and be done with it. One has to show that God would equate suffering and evil, and would regard every form of evil as some kind of suffering.

Maybe that seems like special pleading on the theist's part; maybe it is. But that doesn't invalidate the point. As long as there is one definition of benevolence or one vision of God which contains a loophole explaining your example, that example isn't a contradiction between the existence of some form of God and the world we live in. The atheist proposing the argument from evil, to show a contradiction, thus doesn't have to show that his preferred understanding of benevolence is inconsistent with observation, but that every possible understanding of benevolence consistent with any form of theism is inconsistent with some observed evil. That doesn't just include all definitions and rankings of evil we have thought of; but also all those we have not thought of. And that is a big task; far bigger than most people who have advocated that the argument from evil shows a direct contradiction between the notion of God and the observed world have been able to accomplish, or even tried to accomplish. And it gets even harder when we note the word "consistent" above. Most reputable theists draw their understanding of God from a number of sources, including revelation, observation of the world, and perhaps some metaphysical premises. That observation of the world includes the existence of suffering and other things we might call evil. So the theist understanding of God, and equally the definition of the benevolence associated with the God, will take into account suffering and evil in the world. So any understanding of evil consistent with that particular vision of God will also be consistent with any observed evil. Finding a contradiction in something constructed to be consistent with all the facts and theories you are referring to is a particularly tough task. Thus the logical argument from evil cannot easily be proven.

One response to this is to argue that although we might not to be able to demonstrate a firm contradiction, nonetheless there are cases where it is difficult to justify. Maybe there is a way of getting around the problem in some twisted system of morality that we can't think of, but it is improbable. And that improbability is only magnified when we consider all the different examples of evil in the world, and the difficulty in finding a system with loopholes for all of them. But the problem here is how do we measure the probability. Our knowledge is limited; our knowledge of God's mental processes, if God exists, is minuscule. Thus it could be argued that if we can't see a justification for the particular problem, it is far more likely to be because of our own limitations than because of any real contradiction.

Another response might be, "I view this thing as unjustifiably evil, and if God disagrees with me then He is a monster." Fine, but your opinion doesn't count as fact, neither does it constitute a rigorous proof.

The final approach would be to take a definition of good and evil used by some theists, and use that to find a counter-example to God's benevolence. The atheist might not be able to disprove theism in general, but he can at least disprove that particular vision of theism. But even this is difficult. Ethical theories adopted by theists tend not to be perfectly precise, particularly when comparing the severity of different types of evil. There would still be freedom for the theist to manoeuvre. All the theist has to do is present one plausible theodicy, consistent with his notion of God, and the argument from evil is defeated.

And it is in that spirit that I offer the arguments below. I do not claim that this is the correct answer; that this is how God in practice justifies the various evils in the world. That would be well beyond my abilities. But I do suggest that it is a plausible answer, consistent with both the picture of God developed in the previous chapter, and the ethical theory described in this one. And that, I submit, is sufficient to defeat the argument from evil.

Objection:

(1) Benevolence includes a desire to relieve suffering.
(2) Benevolence is a virtue, and thus part of the character of any good being.
(3) If all Gods are good, then they would not permit suffering to occur if They were able to prevent it and were aware of it. This includes the case of a single good God.
(4) An omnipotent being would be able to prevent suffering.
(5) An omniscient being would be aware of any suffering.
(6) Therefore, if there was one good, omnipotent, omniscient God, there would be no suffering in the world.
(7) There is suffering in the world.
(8) Therefore God is either not good, not omnipotent, not omniscient, has at least one equally powerful evil rival; or God does not exist.
(9) Christianity asserts that God is all of one, good, omnipotent and omniscient.
(10) Therefore Christianity is false.

•

Response:
Note that I have distinguished between the problem of suffering and the problem of evil. I will discuss the problem of evil below.

This objection is in part based on a misunderstanding of omnipotent and omniscient. As defined in the previous chapter, these mean all that is logically possible to know and all that is logically possible to do given the nature of the being that is knowing or doing. It is entirely plausible that it may be logically impossible to prevent all suffering in the world. For

example, a certain type of parasite necessarily causes me suffering. To relieve that suffering for me requires removing the parasite, i.e. causing harm to the parasite. This particular example may be resolved by creating a world where that particular form of parasite does not exist, but if the parasite provides a particularly tasty meal for some other creature, and that creature a particularly tasty meal for myself (or something similar) then we might encounter a situation where it is logically impossible (perhaps baring continual miraculous intervention) to create a world where it is possible to fully satisfy all of these creatures. To deny a creature a joy (such as a tasty meal) is to cause them to suffer in some respect. It is difficult to prove that no such world can exist, but equally it is difficult to prove that such a world can logically exist. It is therefore not completely clear that point 4 in this argument is correct.

The biggest flaw, though, with this argument is contained somewhere within these first two points; in which point it lies depends on how precisely benevolence is defined. It is my view that suffering is not always evil. For example, we can consider the case of physical pain. Pain hurts; there is no denying it, and that is the whole point of it. But pain itself is not an evil; rather it is a signal that some other evil is present in our body. When we injure our leg, the evil is the damage done to the extremity, which prevents it from efficiently fulfilling its purpose. Pain prevents us from causing more damage to the leg; instead it focuses our attention to give it a chance to heal. Would it be for our good if when an injury occurred we felt pleasure, or even nothing? For there is a disease which prevents us from feeling such sensation: leprosy. When a leper places their hand on a hot surface, they do not feel the heat, so the hand burns. When they injure themselves, the body does not respond correctly, and the injury becomes a permanent mark. The leper's body becomes disfigured, weakens, and eventually the attack on the nervous system becomes fatal. That is a world without physical pain. Is it what we want?

When I was a graduate student, I had a problem with my intestines. They became narrowed; food could not pass through my system; I could not eat without suffering intense pain, and gradually withered away. But I ignored the pain; I was too devoted to my work. It was not until my family saw me that they forced me to attend the doctor; and I was fortunate that the doctor was able to find the problem and treat it in time. I nearly died because I ignored the signals my body was sending me. How much worse would it have been if the body had sent me no such signals?

The question then is whether it is logically possible to create a world without a capacity for us to injure ourselves. But surely not, at least not without continual miraculous intervention: we need heat to cook, and in industry. Yet heat also burns us. In some capacities heat is good; in other capacities it can cause evil. One cannot have one without the other. Equally, an axe can be good, if we use it to fell a tree so we can use the wood to construct something we need. Yet if it can fell a tree, it can also be used to injure ourselves. One cannot have one without another, unless there is an inconsistency in that sometimes axes are made of sharp metal and at other times from blunt rubber depending on the use we put them to. This is only logically possible if the world is no longer regular, i.e. it requires an extraordinary act of God. Could we have been made from some impenetrable fully heat-proof substance, incapable of injury, cold or burn? It seems unlikely that any such substance exists, let alone could give rise to a being capable of thought, freedom and reproduction.

Is it logically possible to create a world where 'good' pain, alerting us to some injury, exists, but 'bad' pain, coming from the wear and tear of the body, or those illnesses and injuries we can do nothing about is relieved? Again, this would require the nerves to fire in some instances, but in other instances, identical as far as the nerve endings are concerned if not to ourselves or our doctors. Without continual miraculous intervention, such a world seems to be impossible.

Let us now consider suffering more generally, and ask whether there are any situations where a good being may permit suffering? I think that there are.

(1) There is the case when we have to choose between several evils; say some suffering and some other evil; where there is no logically possible good option. We see such situations in our lives all the time. For example, when Hitler made clear his desire to move his armies into Czechoslovakia, Britain and France were left with a dilemma. Either to let him do so, with the evil that that would cause, or to prevent him in the only way available, by declaring a war that neither were in a position to fight. Whatever they chose, it would have caused suffering either in the short or long term. One can question whether Chamberlain chose the wrong option at the Munich conference; but the real truth was that there was by then no good option available to him. He had to choose between evils. This argument is not, perhaps, by itself enough to solve the issue of why God should permit suffering, since we still need to show that there are situations where it is logically impossible to choose a good; but if such a situation occurred then God would have to, whether through action or inaction, cause someone or something to suffer. Allowing suffering is therefore not necessarily inconsistent with goodness.

(2) There is the issue of discipline. Discipline is a temporary suffering used to bring about good. As it is written,

> It is for discipline that you have to endure. God is treating you as sons. For what son is there whom his father does not discipline? If you are left without discipline, in which all have participated, then you are illegitimate children and not sons. Besides this, we have had earthly fathers who disciplined us and we respected them. Shall we not much more be subject to the Father of spirits and live? For they disciplined us for a short time as it seemed best to them, but he disciplines us for our good, that we may share his holiness. For the moment all discipline seems painful rather than pleasant, but later it yields the peaceful fruit of righteousness to those who have been trained by it[52].

And again,

> Discipline your son, for there is hope; do not set your heart on putting him to death[53].

And again,

> 'You did not do as I had asked. You have never treated Harry as a son. He has known nothing but neglect and often cruelty at your hands. The best that can be said is that he has at least escaped the appalling damage you have inflicted on the unfortunate boy sitting between you.'
>
> Both Aunt Petunia and Uncle Vernon looked around instinctively, as though expecting to see someone other than Dudley squeezed between them.
>
> 'Us – mistreat Dudders? What d'you –?' began Uncle Vernon furiously, but Dumbledore raised his finger for silence (Rowling, 2005, chapter 3).

In this last quote, we see the two evils of which discipline is the golden mean. Firstly, cruelty. The second is over-indulgence; a refusal to discipline, and the result was a fat, gluttonous, under-achieving bully. It was not recorded what happened to Dudley Dursley when he entered the adult world. It is unlikely that he contributed positively to mankind; or if he did it was despite rather than on account of his upbringing.

The difference between discipline and cruelty is that discipline is based on agape love. Discipline is intended to improve; to strengthen character; to make a child realise that it is often better that our wants are unfulfilled. The goal in discipline (which may or may not be achieved in practice) is to do good for that

[52] Hebrews 12:7-11.

[53] Proverbs 19:18.

person. The goal in cruelty is to do harm for the sake of causing harm. It is not agape loving to give vodka to an alcoholic, or drugs to an addict. They will suffer if you (and everyone else) withholds it from them; but in the long term they suffer even more if they receive it. It is not agape loving to give a child the violent computer game he desires, or to share pornographic images with him. In the short term, you are denying him a pleasure; in the long term you are debasing his character making it far harder for him to live a fulfilled life.

Suffering is also necessary to build up our character. Suffering teaches us to appreciate what is truly good. It allows us to see the difference between apparent goods and real goods.

(3) Thirdly (though related), we have the principle of justice. Justice means that everyone receives as they deserve (deserve here relates to what a person is, his character, rather than what they have done). Now it is apparent that if some deserve comparatively good things, then others deserve comparatively bad things. This does not mean simply withholding a pleasure. Evil is a lack of good; any denial of pleasure is a active source of pain for those who are aware of and seeking that pleasure. Therefore, it is good to cause suffering to those who deserve it. A difference between justice and discipline is that discipline is for the purpose of education; while justice is not.

A balance again, has to be made. It is unjust to give pleasure where it is not deserved; it is equally unjust to provide more pain than is deserved. It is not just to cause infinite pain unless the evil is also infinite. Equally, justice is founded in agape love. To reward evil with pleasure is to encourage evil; to reward good with pain is to discourage good; both of these are contrary to the principle of agape love.

Let us try to imagine a world without justice; where evil people are treated in the same way as the good; everyone is left in a state of pleasure. Surely such a world is impossible; for the evil person by definition either directly or indirectly desires that which will cause harm to either others or to himself. Either the desires of the evil person will not be fulfilled, which both causes him pain and is a violation of his freedom, or they are allowed to succeed, in which case there will be pain, and possibly also the denial of somebody else's freedom. It is consistent to have a world where the evil are in pain and the good in pleasure, or (to an extent) the good in pain and the evil in pleasure, or everyone in pain; but not one where both good and evil are in pleasure. Of these, the first is to be preferred. Expanding this beyond the simplified dichotomy between 'evil' and 'good', but acknowledging that there are degrees of goodness and degrees of pleasure and pain, we reach the principle of justice, where each is treated with either pleasure or pain to the degree which they deserve.

If it is unjust to not cause an evil man to suffer, then it is also an evil to not cause him to do so. Therefore causing suffering is not incompatible with goodness; rather, as long as there is some evil, not causing suffering is incompatible with goodness. Justice follows from a desire that good prevail and is exulted; it is therefore a consequence of agape, and to act unjustly is therefore a denial of agape love (and also benevolence which, if defined so that it is a virtue, is an element of agape love). One should also note in this context conjecture 14.6.12.

Does this mean that all suffering is good? By no means! It is unjust to cause suffering to a good person, and to cause too much suffering to anyone. We should be clear: suffering itself is never good; but inaction in the face of evil is even worse. Does this give us the right to punish who we will, since all are evil? With one caveat, no. We have no right to punish someone, because we are just as bad as they are. If we think it right to cause someone pain, then to

be just we would have to cause the same pain to ourselves, and everyone else. Justice should only be administered by a right authority; one with no wrong in themselves, i.e. God. The caveat, of course, is that some degree of justice needs to be administered for the good functioning of society. But again, this should only be administered by a right authority. It is given to governments and courts to see to the punishment of criminals, not vigilantes. Of course, governments and courts will also be accountable for whether they fulfil this task justly or unjustly.

(4) Fourthly, I note that much suffering is caused by humanity abusing its freedom of choice. In this case, one can either allow the suffering or deny the freedom to the people. In either case, an evil must be chosen, and it may be that allowing the suffering is the least evil.

(5) Fifthly, suffering is a continual reminder to us that our world is flawed, and not as it ought to be. It helps us to recognize the depths of evil, including the evil of sin, and thus can help lead us to repentance.

(6) Sixthly, for the Christian, persecution (while never to be sought) is a means of testing and purifying one's faith. The Christian is aware that for themselves the sufferings of this world are only temporary.

(7) Seventhly, suffering allows us to be trained in the exercise of compassion. Once again, no suffering means that people will be unable to exercise compassion, and without exercise the virtue may well atrophy. It is once again a case of choosing between two evils; and it is not clear that God acting to relieve the suffering is worse than allowing those moved by the suffering to build up their virtue.

(8) Suffering need not be permanent. Christian doctrine states that we are given this one brief chance, and then there will come the resurrection and the new creation. If the death of those people who suffered in this way were truly the end of them, then that would be an evil, but that is not the case in Christianity. Equally, in this phase of our life, the state of our bodies is of only secondary concern to the state of our character and our decision of whether to follow God or ourselves. Indeed, such suffering is necessary, since it focuses our mind, and reminds us that the world is not perfect. If we believed the world to be perfect, then we would be less likely to look at the flaws within ourselves, and we would be even more likely to think that we are good enough to live without God. If this world was good enough to satisfy us, would we really desire a better one? Therefore this sense of the objection takes a far too naive view of God's benevolence. God's desire for us is that we should become good, and part of the way in which that is achieved is through such reminders that we are flawed. Is this unfair on those who are caught up in such events? No, what unfairness there is is to the benefit of those who escape them rather than to the harm of those who are involved.

Therefore, causing suffering is not always incompatible with being good (although it often is). Therefore, either benevolence should be defined in such a way that it is not incompatible with suffering, falsifying the first premise of the objection, or benevolence should not be considered a virtue as it would not be good absolutely, falsifying the second premise of the objection. In either case, the objection fails. ∘

Objection:

(1) Agape love includes a desire to vanquish evil.

(2) Agape love is a virtue, and thus part of the character of any good being.

(3) If all Gods are good, then they would not permit evil to occur if They were able to prevent it and were aware of it. This includes the case of a single good God.

(4) An omnipotent being which desired to do so would be able to prevent evil.
(5) An omniscient being would be aware of any evil.
(6) Therefore, if there was one good, omnipotent, omniscient God, there would be no evil in the world.
(7) There is evil in the world.
(8) Therefore God is either not good, not omnipotent, not omniscient, has at least one equally powerful evil rival; or God does not exist.
(9) Christianity asserts that God is all of existent, one, good, omnipotent and omniscient.
(10) Therefore Christianity is false.

•

Response:
The problem of evil, of course, is not evidence for atheism as such. The obvious solution to it is the belief that there are two equally powerful forces in the world, one good, one evil. This leads to a world which easily explains why the world appears to be partly good and partly evil. Such a system provides the basis for the religion of Manichaeism, a rival of Christianity in the attempt to replace Greco-Roman-Mesopotamian polytheism with something more philosophically rigorous. Of course, Christianity won that battle; the philosophical arguments more strongly point to a single God rather than the dual system of the Manichees, and, of course, the empirical (observational) evidence for Christianity is far stronger than any of its rivals. However, I note, before discussing my main argument, that Christianity does outline a battle between good and evil spiritual powers. The difference with Manichaeism is that where that religion sees the evil power as existing in itself; Christianity sees it, like everything else baring the single necessary self-existent being, ultimately being caused by and dependent on another. It is not my intention to focus on angels and demons in this work (since evidence for their existence arises from revealed rather than natural knowledge), but some mention ought to be made here. Just as with God, we should immediately abandon any cultural pictures we have in our mind of the devil or an angel, even those promoted by the Church. Such images only serve to distract us from the real thing. The Biblical witness, however, is clear. According to it (and how seriously one takes this depends on how seriously one takes Christianity or its related religions), there are immaterial minds other than God; they do have a certain limited degree of power to influence this world; some of them use this power for evil, and they are in practice responsible for a great deal of evil that we see around us. Such beings are not God; they are not immutable, and thus have the ability to actualise some part of their form which is not suitable to fulfil their final cause. The same choice given to us has been given to them; like us some have chosen to be faithful to God and are therefore protected by God from straying; others have chosen rebellion. Those who choose rebellion have caused a great deal of damage to our race.

It is often difficult for us to tell which are good and which are evil, and many have been deceived following an evil spirit whom they believe to be good. "For even Satan disguises himself as an angel of light[54]." To distinguish one from another requires external knowledge outside merely angelic or demonic revelation to test whether or not the revelation is true or false. For Christians, this knowledge comes partly from philosophical arguments, but mostly by looking at the person of Jesus; for Jesus (if Christianity is true) was proved by great works to not be a mere angel or a mere man guided by an angel but God Himself dwelling among us[55].

Of course, such an explanation has problems; it can be no more than a partial explanation for the existence of evil. Firstly, I do not know of any philosophical argument for

[54]2 Corinthians 11:14.
[55]Hebrews 1:1-2, John 10:24-25,30-38.

the existence of such creatures. One comes to them indirectly, first by accepting the direct evidence for Christianity (or some other religion that testifies to their existence); and then necessarily accepting the testimony of Christianity (or that other religion). That is not to say that such being's existence is inconsistent with reasoned arguments, merely that (unlike God) it cannot be proved from reason coupled with basic observations about the world. Secondly, it is not an excuse for our own evil behaviour. The excuse "The devil made me do it" will not carry any weight on the day of judgement. We may have been influenced, directly or indirectly, by such beings; but the final choice is still ours, and the final responsibility is still ours. Thirdly, it only offers a partial solution - we ourselves are not innocent bystanders, but actively responsible for most of the evil we see around us. We cannot blame everything on Satan (if he exists), because Satan is not responsible for all the evil in the world. Fourthly, and most importantly, it does not resolve the original problem. I mention the issue because Christianity teaches that much – not all - of the evil we see in the world is caused directly or indirectly by the actions of the devil. But this simply pushes the question back one level: where did the devil get his evil from? We cannot use this excuse indefinitely.

Now, before discussing why the objection above fails, I ought to remind the reader that evil is a lack, or distortion of something good. Evil does not exist in the same way that a good exists; rather evil can be thought of as a lack of (true) existence. Part of the form of a being is in alignment with the being's natural purposes; most of it is not. To be evil means that the actual state of the being is described by that part of the form not aligned with the final cause; the being lacks some capacity or attribute that it ought to have.

Secondly, the definition of omnipotence given in the previous chapter was not that an omnipotent being can do anything imaginable, but only anything logically consistent and consistent with its nature. So, for example, suppose that it is essential to Man's nature that he has freedom of will; this means that his desires are respected. Suppose that this Man is offered a choice; either he can live in a world with frequent breaks from the usual natural order, or one where he can live seemingly independent of God but in a world where everything proceeds according to natural patterns. The first of these is the path of humility - Man acknowledging his dependence on God. The second is the path of pride – Man believing that he can exist well seemingly independently of God. Suppose that Man then chooses pride. This, of course (for example) entails natural disasters. So what is God to do? On one hand, he violates the free will of his creation - an evil: Man is not able to fulfil his natural ends. On the other, we have the bodily suffering which (except as a requirement of justice) is also an evil. Can God procure a world with the benefits of miraculous intervention to prevent natural intervention but man able to live in disobedience? Firstly this will be a denial of justice. Secondly, mankind's disobedience is precisely in his rejection of God's protection from harm, which would include the harms caused by mankind's own bad choices. To be consistent, God would not only have to suppress nature's ability to cause harm, but also to repress against our wishes our own internal desires and free choices, since we are part of nature. But this will deny us freedom. The only other option I can think of is to not create beings with free will in the first place – but this too is an evil, since it denies God the opportunity to display his love for creatures such as ourselves: every lover needs a beloved. Thus God is left with a choice between evils; and if He should choose the lesser evil, then that is still an evil. Thus there will be evil in the world of one sort or another unless all existent Men choose to humble themselves before God. If this is correct, then some evil cannot be necessarily avoided. If it is logically impossible that there is no evil, then the existence of evil is not inconsistent with the existence of an omnipotent God.

If evil cannot be avoided, then the question turns to how it can be minimised. This is in principle a difficult question (as we do not, and possibly cannot in practice, know everything we need to understand to solve it). Nonetheless, I can suggest a partial solution. Clearly, all the evils are avoided if Man freely chooses goodness. For this choice to be informed, then it must be possible for him to attain knowledge of good and evil, including their nature and

consequences. In practice, this means that can be two phases to the life of Man; one brief moment where he has to make this informed choice, and consequently will experience evil in order to understand it; the second (far longer) when he lives with the consequences of that choice after a general resurrection (i.e. the recombination of Man's form with matter). In the first (far briefer) phase of his existence, the means through which Man can also attain the life of goodness should be revealed and made clear; the path made possible, and enough information made available to inform the choice, but not so much that it overwhelms Man's experiences. The evil of the first period is over a far shorter time than the good that follows (which may be essentially infinite). Obviously, Man is not a single entity; individual men and women must make their own choices, and some would choose the path of humility, and others the path of pride.

This story is, of course, precisely what is envisaged in orthodox Christianity; I have merely skipped over most of the details. The second phase, of course, follows the creation of the new heavens and the new earth ('heaven' in true Christianity is not the vision of angels, clouds, harps, or some disembodied spiritual realm which is sometimes imagined in popular culture, but a remaking and perfecting of our current world). One detail I have skipped over in this brief review is that of hell: the place where those who chose pride will dwell without 'interference' from God. This will, however, be a place of torment, since man's natural state tends towards misery; the general grace we receive today that allows us to experience some happiness will be removed. The torment is just, for it is not solely based on the deeds of men during this time, but their character, which, without the working of God, can only deteriorate. Since they will never reach the standard required for heaven, since that is impossible without God's help which they refused, they will be left in continual wailing and gnashing of teeth.

Furthermore, there is the additional problem that in destroying evil, one may also pluck out much which is good or potentially good. For example, consider the following parable,

> He put another parable before them, saying, "The kingdom of heaven may be compared to a man who sowed good seed in his field, but while his men were sleeping, his enemy came and sowed weeds among the wheat and went away. So when the plants came up and bore grain, then the weeds appeared also. And the servants of the master of the house came and said to him, 'Master, did you not sow good seed in your field? How then does it have weeds?' He said to them, 'An enemy has done this.' So the servants said to him, 'Then do you want us to go and gather them?' But he said, 'No, lest in gathering the weeds you root up the wheat along with them. Let both grow together until the harvest, and at harvest time I will tell the reapers, Gather the weeds first and bind them in bundles to be burned, but gather the wheat into my barn.' "
>
> Then he left the crowds and went into the house. And his disciples came to him, saying, "Explain to us the parable of the weeds of the field." He answered, "The one who sows the good seed is the Son of Man. The field is the world, and the good seed is the children of the kingdom. The weeds are the sons of the evil one, and the enemy who sowed them is the devil. The harvest is the close of the age, and the reapers are angels. Just as the weeds are gathered and burned with fire, so will it be at the close of the age. The Son of Man will send his angels, and they will gather out of his kingdom all causes of sin and all law-breakers, and throw them into the fiery furnace. In that place there will be weeping and gnashing of teeth. Then the righteous will shine like the sun in the kingdom of their Father. He who has ears, let him hear[56].

[56]Matthew 13:24-29,36-43.

The weeds are, of course, of a nature such that they are indistinguishable (at least to us) to the wheat until the crop has matured. The argument, I think, is the following. Suppose that at one particular moment in time, God were to intervene and banish all evil from the world. This would, of course, include banishing all those people who desire to be evil, since one cannot make a free being good against their will without violating their freedom: a contradiction. But this will reduce the number remaining who desire good; firstly because some of those presently bound by pride may subsequently change, and secondly on account of those who are driven to embrace agape in response to manifest evil. Thus, in the process of plucking up the weeds, it is inevitable that some of what will become wheat would also be removed. Of course, if God were to repeatedly and obviously step in, then many might seemingly join Him for the wrong reasons, out of fear rather than agape. To maximise the number of people who might receive mercy, God is patient, and lets things play out until no more people would be suitable for receiving mercy.

Another way of looking at the problem is to ask whether a being motivated primarily by *agape* should seek to minimise evil, or maximise good, or perhaps something else. One question is how we should count evil. Is evil to be counted as something negative, which can cancel out good in the measure we use to compute the total amount of good in the world, or merely as nothing at all? Is the scale (if one can excuse the crudity of the image) from -1 to 1, with -1 purely evil, and 1 purely good, or from 0 to 1, with 1 perfectly good, and 0 a complete absence of good (and thus what we call pure evil)? Given the idea that evil is an absence of goodness rather than a negation of it, then the latter scale should be preferred. Of course, this picture is overly simplistic: even if we could assign the degree of goodness to a numerical scale (which I doubt), there is no reason why each type of being should have the same maximum value. For example, some (such as those which possess a free will) would be able to obtain greater goodness than others.

For example, suppose that some deity had a choice between two different worlds: the first is one with one being as close to being good as a non-divine creature can be. The second option is one where there are three half-good creatures of the same type. Clearly, if the aim is to minimise evil, or maximise good minus evil, then the first option should be chosen. If, on the other hand, the aim is to maximise good, then three halves is better than one. Indeed, if God's aim was to minimise evil, or the capacity for evil, then He would not create anything, because only God is by nature perfectly good. Now, obviously this argument is somewhat simplistic. Why not create a world with three almost perfectly good beings - surely that beats the world with three half-good beings?

The first reply to this is that all existent beings are to a certain extent good, because existence in itself is a good. The creation of any being, no matter how much goodness (or lack thereof) they possess, therefore will still increase goodness. The second part of the answer is to suggest that beings with a free will almost certainly count as more in the equation than an inanimate object. Almost nobody chooses to be bad: usually the problem is that people pursue one particular good too much, and so move away from the golden mean of virtue into some sort of vice. For example, it is good when a free will aligns itself towards some sort of goodness. If an inanimate particle follows its final causes, then that adds nothing to the total goodness of the universe, since ultimately it is only a reflection of the goodness of God. However, if a rational agent, acting (from one point of view) independently of God chooses to align its will to a certain extent with its natural purposes, even if not completely, then that adds to the goodness in the universe. Thus, from this perspective, it is in God's interest to create agents which can freely choose to desire good or not.

It is also evident that if one creates rational agents which are free to either choose good or to choose to avoid good, then at least some of them will choose to align their wills at least partially against goodness. For if they were all manufactured in such a way that all of them would automatically happen to choose good, then that would mean that they were not truly free in the first place. It is also evident the the spot of perfect goodness is only a tiny

spot on the manifold covering all possible virtues and vices: there is only one small region of perfect virtue, and an infinite number of ways of falling into vice to a greater or lesser extent. For example, let us consider a two dimensional example, which is easy to visualise. Take a black piece of paper. The left side of this paper represents the vice of cowardice, the right the vice of recklessness. The top of this paper represents the vice of apathy; the bottom the vice of the over-indulgence of others. There is one microscopic dot, somewhere in the middle, which represents the position of virtue: courage and benevolence. Now a free will may choose to align itself anywhere on the paper: if it is forced to choose one spot, then it is not truly free to choose good, and unless it is free the choice will not contribute additionally to the total sum of goodness as discussed above, but only be a manifestation of God's goodness. If so, then the chances of it happening to pick the one spot that represents perfect virtue is very remote. But, unless it chooses a spot on the edge of the paper, it will always pick a place which represents some degree of goodness, if not perfect goodness. Thus each new individual will increase the total amount of goodness, even if they will not be perfectly good and thus contain a certain degree of evil. In this way, the world with the maximum amount of goodness (and the maximum number of people) will still contain a great deal of imperfection.

Each point on the manifold represents one possible example of what our character could be (recall that goodness is a matter of who we are primarily rather than what we do). Character depends on virtue and vice, which are related to our will, since they indicate how much our desire is aligned with our natural purposes. It is thus important to distinguish between two different aspects of our will: our conscious will (which in principle, if not always in practice, we can control), and our unconscious will (which represents who we are, our virtue and vices). For example, someone can desire to be good while still in practice falling short of those desires (though weakness or negligence or some other reason), or earnestly intend not to get angry, but still in practice losing their temper. But being good means having our will aligned with our natural purposes. To avoid a paradox when we talk about failing in our desire to desire to follow virtue, we must accept that each of the words 'desire' in that statement is intended in a different sense: one refers to our conscious will, and the other to our unconscious will. Our unconscious will is influenced but not determined by our conscious will: there are other factors that affect it, such as our inborn nature, character, and so on. While it can be trained, such training takes time, effort, and a great deal of determination. Similarly, our conscious will follows our intellect, which is in practice in part influenced by our character. Having a free will means that both our conscious and unconscious wills are not fully determined by our nature. Thus there are two points on this chart: one represents what we are, and the other what we desire to be. The second of these is (for the one who is mature enough to be able to control their thoughts) wholly a matter of choice; the first is only partly a matter of our choice. The truly good person will have both of these wills perfectly aligned with virtue.

Thus, if these beings were to 'randomly'[57] align their wills (which is surely what has to be meant by having a free will: there has to be some distribution of possible choices), only a tiny fraction of them, if any, would choose (both consciously and unconsciously) perfect virtue. There would always be some evil: indeed, there would always be a great deal of evil. Now, of course, their conscious desire would be directed by their intellect, and, if used properly, their intellect would direct them towards goodness. But this does not change that their wills will be distributed seemingly at random; it merely changes the distribution of where the free wills align themselves: it will not be uniform across the whole range of virtue and vice, but peaked at or near the point of virtue. But knowledge is imperfect, indeed it

[57]It would not be random as such, since each individual person would have their reasons for consciously choosing to desire what they do (as much as their unconscious desires are a matter of their choice). However, for a large number of people, the total distribution of these desires would be indistinguishable from a random distribution.

is far from perfection (except in God), so even if we consciously desire to be good, we will still miss the mark because we are not fully certain where the mark is. This peak will not be sharp enough to ensure that the most of the rational agents naturally choose a path of perfect virtue. So small is the target, that hardly any of them (if any) will do so. But each of them, by coming close, will still contribute to goodness, since none of them would be wholly evil. Thus there is still more goodness in creating this world of imperfect beings than there would be of creating a world with no beings at all, or only beings with no freedom. So if this picture is valid (and I have made a number of assumptions which could be challenged), then not only is evil possible in a world created by God to maximise goodness, it is pretty much inevitable that there would be evil, indeed almost all of his rational creatures would be evil to at least a limited degree. This is, of course, what we see in the world around us.

To return to my crude numerical analogy, suppose that a rational being without free will and the ability to choose between good and evil can score a maximum of two points of goodness. However, to use a free will well is an additional good. A being with free will has a greater capacity for goodness, since their free will is an additional good. So let us suppose that such beings can score a maximum of three points of goodness. The beings without free will in this regard will each score two points. Those with free will will score less than three, since all of them will fall short of the ideal in some respect, and evil is defined as any falling short of the ideal. However, they might still average more than the two points per being that God would have achieved if he did not create them with free will. While this picture is crude, and should not be taken too seriously, it still illustrates that the world with the largest amount of goodness might well contain some evil, while a world no evil might have less goodness. If a benevolent being is one who desires goodness, then He might well allow a world with some evil in order to have more good. Of course, once the free will beings become so irredeemably evil that their score drops below two, they might expect to face some rather unpleasant judgement. But above that point, God would put up with some evil in order to have more good.

Now I should emphasise that this picture, which I have tentatively presented, differs from orthodox Christianity in at least one important respect: Christianity teaches that our wills (especially our unconscious wills) are not by nature perfectly free: we are slaves to sin. In the picture, there is a vanishingly small, but still non-zero, chance that a man would by chance happen to have his will aligned with the point of perfect virtue. In Christianity, this chance is zero: we have an inherited corruption which means that we cannot be good enough to satisfy God by our own efforts; the best that we can do is consciously desire that God acts to make us perfect. This, however, only presents a minor modification to the overall picture. This means that everybody has to consciously choose goodness (and thus implicitly choose God); there is nobody who by nature happens to be good, whose goodness would thus be merely a manifestation of God's goodness rather than in addition to it.

The key error in the objection is point 4. If one desires to create a being which has the possibility to choose between good and evil, then it is logically impossible to then prevent it from choosing to be evil if it so desires. To deny such a being the act of becoming what it chooses it to be is to deny it part of its natural purposes; an evil in itself. One can, of course, create a universe where no such beings capable of such a choice exist. But this universe is not as good as one in which beings with free choice do exist; because to deny such a universe is to deny the good of existence to that type of being[58]. In the same way,

[58]Note the phrase 'type of being'. One can, of course, imagine an animal with (for example) the body of a cow and the mind of a man. Let us call such a creature a cowman. Such a being does not (as far as we know) exist (there may, of course, be cowmen on a planet in a distant solar system, but whether there are or not is irrelevant to the argument here). Is, if the argument that a lack of rational animals is an evil because it denies the good of existence to such a genus, does the same argument mean that the lack of the species of cowmen is also an evil? Not as such; the key issue is whether or not a rational animal should exist. The form of the material (animal) body of the rational animal is relatively unimportant; one must pick one or a few choices, such as ourselves, or cowmen, or an infinite number of other beings. Clearly, the number of such

to create a universe which is incapable of any significant change (also a way of eliminating the direct evil of things being unable to fulfil their natural purposes) is an evil in itself, because it denies existence to beings capable of change. It is therefore logically impossible to create a universe without evil. Either the evil will manifest itself in beings which are evil in themselves (not aligned to their final cause), or it is evil because certain types of being which logically could exist do not exist. The existence of evil, then, is not inconsistent with both omnipotence and agape love, once omnipotence and the nature of evil are understood correctly.

A partial solution to this dilemma, and I think the best that can be achieved, is to limit this evil as much as possible. This can be done by making the evil temporary. We would create a universe which allows evil for a short time, before fixing it as good. Evil therefore exists, but is not permanent. This is done by giving rational animals a choice: to live in a universe where they are constrained to remain good, or not. This is not against the freedom of choice of such creatures, because they have freely chosen to be so constrained. Those who do not freely choose good continue to exist (for denying them existence will be an evil) but, for the sake of the good of justice, in a state of punishment (to leave them otherwise would be an evil[59]), But this, of course, is precisely the universe that Christianity teaches us we are living in.

Now this argument tells us why rational agents such as ourselves can be evil. This accounts for much of the evil we see around us; terrorist attacks, genocide (including abortion), rape, theft, slavery and so on. To forbid such things is to deny people their freedom. What about the rest of the world? For example, earthquakes, volcanoes, typhoons, cancer, and, of course, the many small things which go wrong without being anybody's fault. In the previous response, I related these to justice; such evils occur to us because we (as a species) deserve it (I do not mean that those who suffer most deserve it the most: only that we all deserve suffering, but some by grace are spared it to a certain extent). The Christian answer, of course, goes further than this (though incorporates this; see Luke 13:1-5). Whether one takes Genesis 2-3 literally or largely metaphorically, the account states that we were offered the option of living in a world without such things. Our ancestors rejected it. We told God that we would rather live in a world without His direct day-to-day involvement, a world where we seemingly rule, than in the world that God designed for us. Our ancestors said to God 'We don't need You', and God replied 'OK; see how well you fare on your own,' as agape compelled Him to respect our freedom of choice. And, of course, the bulk of humanity continues to say to God 'We don't need You' to this day. Earthquakes occur because we do not want God to intervene to prevent them; or, more accurately, we do want to do that which enables God to intervene while simultaneously respecting our freedom: to surrender our pride; to cease our rebellion.

Why then does God not intervene to put everything to right? To do so too early would be to limit the number of people who get the chance to choose Him; for an intervention of that scale would mean the end of the world as we know it, and free choice as we know it today would be impossible afterwards given that putting everything right includes fixing us according to our choices. To do so too late will allow the amount of evil to multiply. We therefore can suppose that God will intervene at just the right moment. It is difficult for

---

beings which can exist in harmony on a single planet is limited. Therefore, it is impossible to substantiate every possible kind of rational animal on a given world. The number of different species of rational animal will therefore be limited. It is difficult to judge whether this can be achieved through one, or two, or three such species. But to achieve the good of having a rational animal exist, only one of those kinds needs to exist. The good is that at least one kind of rational animal exists. On this planet, that happens to be us.

[59]I am not saying that I rejoice in the suffering of the damned. I would much rather that those people did not suffer; but logically the only way for that to be avoided is through repentance and sanctification, since to remain in a state of evil is to suffer. If they choose not to repent, then I weep for that choice, but acknowledge that logically justice then leaves no choice but to accept that ultimately the punishment is a good.

us, with our limited knowledge, to say when this will be. One possibility would be if evil seemingly triumphs; a man of lawlessness arises who scoffs at all virtue, and establishes a system of government which makes seeking virtue and seeking God effectively impossible. Such systems have appeared incompletely on a small scale from time to time (for example in Nazi Germany; the Stalinist Soviet Union, and today in places such as North Korea and those parts of the Islamic world subject to the most faithful forms of Sharia); if such a system appears on a global scale, then God would intervene. There would be no more opportunity for righteousness. There would be no point in waiting any longer.

Finally, I note that to have an objective understanding of 'good' and 'evil', which the objection assumes, requires that there is some final causality; and as we saw before, if there is final causality, then it is difficult to avoid having a personal God. Thus the existence of evil (if there is indeed evil) is an argument in favour of the existence of God. (We have the existence of evil $\Leftrightarrow$ final causality; final causality $\Rightarrow$ God). The question then becomes whether the particular understanding of God we hold is consistent with the existence of evil: some are, and some aren't. This objection is not a good reason to adopt atheism or agnosticism. ○

Objection:
The response above depends on a general resurrection of the dead. However, 'when we die, we rot.' Resurrection goes against the laws of thermodynamics, and is therefore impossible. Additionally, the matter from which we are made is spread out, so some of the particular atoms in my body would have been part of another man before me. But this is impossible if we are both to be reconstituted from the same matter. Thus the argument above cannot be valid. •

Response:
The laws of physics are an outworking of God's actions in sustaining the world. They describe how God usually operates in the absence of particular purposes. Our forms are maintained in God's intellect, and recombining these forms with matter is certainly not beyond the power of God. Thus this objection, being based on a false conception of physical law, and therefore false. ○

Objection:

(1) There exists some evil, E, in the world.
(2) Theism, T, is not intrinsically much more probable than some metaphysics, I, (such as atheistic naturalism) where the universe operates indifferently to good and evil.
(3) The probability of E given I is much greater than the probability of E given T, $P(E \setminus I) \gg P(E \setminus T)$.
(4) Therefore, by Bayes' theorem, I is much more likely than T.

Discussion:
By "intrinsically improbable", it is meant before we look at any evidence. In particular, a simple and more general proposition is seen as being more intrinsically probable than one which is complex and specific. So, for example, the proposition "When I go out for a walk this afternoon, I will pass a man" is necessarily more likely than "When I go out for a walk this afternoon, I will pass a man with a wooden leg, eye-patch, and a parrot on his shoulder who answers to the name of Charlie." The reason for this is because the probability that I pass any man includes someone like that, but also a vast number of other possibilities. It is

therefore much more intrinsically likely that I pass a man than I pass that particular man. This can be changed by evidence; for example if I happen to be on a pirate ship it is more likely that I will encounter our wooden legged man than otherwise, but that is additional evidence; it is not part of the intrinsic probability.

This is relevant in this case because the assertion is that although we could invent some form of theism in which E is a justifiable evil, this is very specific and thus has a low intrinsic probability compared to theism in general, when compared to I in general. In the vast majority of the possible versions of theism, E is unjustifiable. Thus if you want to avoid the problem of evil, you instead create the problem that your preferred very specific version of theism is highly intrinsically improbable.

In more concrete terms, suppose there are numerous different ways in which we can construct T: T_1, T_2, T_3 and so on. A similar consideration will apply to I, but it is presumably less relevant because the probability of E given I will not vary much from one construction of I to another. So let us suppose that in T_1, E can be justified in the particular construction of what counts as good and evil used by whatever God is implied by T_1, perhaps because from this God's perspective, not E is an even greater evil than E. And this is true for every possible evil E in T_1 alone; all the other vast number of variants of T fall short for at least some evil (this assumption can be relaxed without affecting the argument, but I will keep it to make the presentation simpler). Then we write,

$$P(T \setminus X) = P(T_1 \setminus X) + P(T_2 \setminus X) + P(T_3 \setminus X) + \ldots,$$

where X is whatever neutral background information (prior to any observational evidence) is universally agreed on. Intrinsically, there is no reason to prefer one variant of theism over any other, so if there are N self-consistent constructions of theism, then we can write $P(T_1 \setminus X) = P(T \setminus X)/N$, which is going to be a small number. But what we are interested in is the probability where T is the condition, not the event, $P(E \setminus TX)$. We use Bayes' theorem to write,

$$\begin{aligned} P(E \setminus TX) =& P(T \setminus EX)\frac{P(E \setminus X)}{P(T \setminus X)} \\ =& (P(T_1 \setminus EX) + P(T_2 \setminus EX) + P(T_3 \setminus EX) + \ldots)\frac{P(E \setminus X)}{P(T \setminus X)} \\ =& (P(E \setminus T_1X)\frac{P(T_1 \setminus X)}{P(T \setminus X)} + \ldots) \end{aligned}$$

$P(E \setminus T_1X)$ is relatively large, but it has a weight of the order of $1/N$. There will be N other terms in the series, of course, but here $P(E \setminus T_2X)$ (for example) will be small. So the fact that there is one variant of theism in which E is justifiable does little to make the overall probability $P(E \setminus TX)$ large; it will be much smaller than $P(I \setminus TX)$.[60] □

Response:

Firstly, there is a difficulty in assigning numerical values to any of these probabilities. To say that something has a low probability does not mean that it is untrue; it means that given our assumptions, there is a low reason for considering it true given the numerous outcomes. If we made different assumptions, we would get different results. In particular, ideas about God tend to be constructed by looking at the evidence, including the evidence from evil, rather than writing down every possibility and then holding a raffle to decide which one is true. The assumption that each variant of theism should be given equal weight fails because the way we construct models of theism immediately prefers some over others.

[60] Arguments of this sort have been presented by (Draper) and (Rowe, 1996) in response to the failure to prove the logical problem of evil, namely to show that the existence of evil is inconsistent with every possible model of theism. By 'Model of theism' I include an ethical theory, which includes identifying good and evil, and ranking which evils are worse than others.

Secondly, the argument only looks at one particular type of evidence. It might be that we have independent reasons to prefer T over I, or even T_1 over the other forms of T. Once we include this evidence, then the probability will shift, and it might be that T_1 is no longer so unreasonable.

Thirdly, we should question whether it is appropriate to invoke probability at all to decide between theism and indifference. Probability, after all, is always conditional on its premises, so the values you get are entirely dependent on the assumptions put into the model. In that sense, the probabilities are arbitrary. At the end of the day, with any probabilistic model, we want to compare it against some empirical test. That means that we compare the calculated probability distribution against some measured frequencies. But what does it mean to have a measured frequency distribution of philosophical theories? In the absence of a frequency distribution to compare against, or some symmetry that allows us to establish a probability distribution according following the classical interpretation, any probability between 0 and 1 just means it might be true or it might not given the axioms used to calculate the probability. Which doesn't really help us determine the truth of a particular theory. All this argument is doing is saying that there are a lot more models of atheism consistent with the existence of evil than there are models of theism. But that doesn't make that model of theism less likely to be true. It only makes it outnumbered.

Fourthly, the argument asks the wrong question. We are not interested in comparing T in general against I in general. Nobody worships a generic theistic God; they worship one specific version of the Christian God, or one specific version of the Islamic God, or something similar. The Bible and the Qu'ran and other religious texts are well aware of the existence of evil in the world, and provide their own way of incorporating it into their understanding of God. Maybe not a complete answer to the problem, but the seeds of a complete and robust answer will be there. So these religions will have an intrinsic bias towards the God of T_1 in comparison to those theologies and theories of ethics where the existence of evil is a major problem. We want to know which specific model of how the universe operates is most consistent with the evidence. We should therefore be taking the best specific model of T (perhaps T_1, although 'best' includes other evidence than just that from the problem of evil) against the best specific model of I, and then ask what what is $P(T_1 \setminus E)$ in comparison to $P(I_1 \setminus E)$. Our question is "Is T_1 more consistent with reality than I_1?" rather than asking about theism in general. And when you start focusing on individual models of both theism and an indifferent universe, then a key step in the argument falls apart. Why should $P(T_1 \setminus X)$ (and presumably $P(I_1 \setminus X)$) depend on how many theistic (or atheistic) models we have? For example, if we have a thousand theist models and a hundred indifferent models, then this construction would make I_1 intrinsically ten times as likely as T_1. Yet, by the initial assumption, I_1 and T_1 both equally well explain the evidence X. Why should we give them different initial weights just because we have constructed more theistic models than we have indifferent models? And if we do compare I against T_1, then that is intrinsically unfair, because we are comparing a whole set of models on one side against a single model on the other.

Why arbitrarily only group all the theist models together? For example, if we enlarged T to T' which contains every possible model of theism and the most popular few models of atheism A_1, A_2, and A_3, and I contains every possible model of deism, then we could have a similar argument. With three atheist models included, each model within T' is then weighted by $1/(N+3)$. Certainly, there are now four models within T' which are consistent with the existence of evil rather than just the one there was within T. But these will still be weighted down by all the theistic models which are inconsistent with the existence of evil. The end result would be the conclusion that deism is far more likely than theism and these atheist models combined, and therefore those atheist models also fall to the problem of evil. Of course, this argument is absurd. Surely there is no reason to group A_1 with T_2, as they are very different models of the universe, one being atheist and the other theist?

But T_1 might be just as distinct from T_3 as A_1 is from T_2. T_1 might share many features in common with A_1, just disagreeing about the existence of God, but agree with T_3 on the existence of God but disagree with it in every other respect. So if one can argue that A_1 shouldn't be placed into T' because it is different from all the other models, one could also make the same argument for T_1. If we are justified in separating atheistic models from the theistic models when we construct our sets of models, why are we not justified in separating those models of theism which are consistent with the existence of evil from those which have a problem with it. Make T just the set of theologies which allow for evil, and then try reproducing the argument? Grouping the models which can cope with with existence of evil together with those which aren't just arbitrarily biases the argument against theism. Why should we discount model T_1 just because model T_2 falls to the problem of evil? ○

•

CONJECTURE 14.7.1. *A benevolent, omnipotent God would not remain inactive in the face of evil.*

Proof:

(1) A benevolent being desires to increase good.
(2) Evil is a privation of good.
(3) Therefore, where there is evil, goodness can in principle be increased by transforming the evil elements of that substance.
(4) A benevolent being would only be inactive in the face of evil if He was incapable of changing things.
(5) An omnipotent being is capable of changing evil to good.
(6) Therefore God would be active when faced with evil in trying to reduce that evil as much as possible.

■

Discussion:
This is not the same problem as why evil exists[61]. Evil exists, we have seen, because it is impossible to have truly free wills where there is no evil; and the existence of beings with free wills, even if they are abused, is nonetheless a greater good than if those beings did not exist. However, that does not mean that evil cannot be minimised if those wills cease desiring it. Now, to cease desiring evil, they need to understand the nature of evil, and especially their own evil, and also trust in God to heal them of the evil (since people cannot achieve that themselves, but need God's activity). It is not evil to correct a character that wants to be corrected; it is not evil to provide information so that an honest and informed choice can be made. The desires of wills can be changed from seeking evil to seeking God. A will can not be forced, but it can be persuaded. Thus we expect God to be active in revealing and persuading people to choose goodness, as well as providing the means by which those who seek goodness can be transformed and maintained in a good state while still satisfying the requirements of justice. Too much activity by God will in effect force people to be good, not out of a desire of goodness, but from a fear of God. Too little will not have the desired impact. Thus we expect there to be some optimal amount of intervention by God, which correctly balances both requirements: providing the needed information to those who seek it, and not so obvious that those who do not desire God's intervention will be overwhelmed.

[61] I use the word 'existence' here and elsewhere in reference to evil in a metaphorical sense; of course, in reality, evil is not existence but a lack of existence.

Thus we expect God to act directly in the world, to inform humanity not only about what it ought to be, but also about God's nature, and how God is willing to help those willing to accept it to achieve their nature; plus to perform any acts required for that transformation to occur. Thus we expect that there have been some historical examples of obviously divine intervention in human history.

Combining this conjecture with the observation that there is evil in the world, we note that we can expect God to be active in the world. This rules out those religions which leave God as an inactive bystander to the world, most especially extreme forms of deism. □

Objection:
I can't believe in your God because It is immoral. It hates gays/condones slavery/is racist/hates women/encourages genocide/some other expression of immorality. •

Response:
This objection is frequently raised, but obviously of little substance. One is comparing an universal, objective and certain morality, against one's own subjective and uncertain ethics, and saying that the subjective beats the objective. If God exists, and is of a certain character, then as the one who establishes the natural purposes and forms of everything in the world, is responsible for framing the objective picture of what is moral and what is immoral; and thus the true theory of ethics cannot contradict the character and nature of God. If God hates homosexuals, then it is right to hate homosexuals, and if we do not hate homosexuals then we are in the wrong. This objection is not based on reason, it is based on an inability to accept that we (i.e. our premises or line or argument from those premises) might be wrong; a form of pride.

Now, I should add that I am firmly opposed to the notion that the God of Christianity hates homosexuals (rather, He displays His agape for them by opposing those actions which are harmful or degrading to them) or women, or approves of chattel slavery, or judges certain races of men as superior to others. These are all misrepresentations of Christianity, or based

on a lack of historical knowledge[62], most commonly based on misunderstandings about what it means to hate.

The charge of genocide is more complex. There is a great deal of violence recorded in the Bible, especially in the Old Testament. However, it is important to realise that just because people in the Bible, even people generally approved by God, act in a certain way, it does not mean that God approves of that action. For example, Lot, who is generally regarded as righteous, committed incest with his daughters; that does not mean that God approves of incest, far from it. David, described in the Bible as one of the best men, had many wives and concubines, and committed adultery; yet God demands monogamy. All this means is that even David fell short of God's standard. It is a reminder to us of the claim made repeatedly throughout the Bible, and which is central to Christianity, that none of us, not even the best (and we can debate whether Lot and David were among the best) are capable of meeting God's standard; everyone is capable and indeed prone to falling away from God's standard. Some, like David, repent and come back. Others, like Lot and Esau, do not, and sell their birthright for nothing. Most of the wars in the Old Testament; most of the violence committed by the Biblical heroes, are either in self defence or against the will of God, and thus says nothing against the character of God.

[62]For example, the situation regarding Christianity and slavery is far more complex and nuanced than many critical accounts I have read seem to realise. Every society in the ancient world was built upon slavery. Every society regarded it as normal. Two religious movements or philosophies fought for the slave: Buddhism in the East, and Christianity in the West. The Buddhists had a few temporary victories in India and China, but these were invariably reversed after just a few years. Christian doctrine immediately undermined the chief argument for slavery – that slaves themselves were an inferior breed of men (e.g. Aristotle, *Politics*, 1.5) – though in the early days some conceded the practice to be an evil but necessary economic reality, since, it was argued (rightly, or more likely wrongly), the complete economic collapse caused by abolition would be a greater evil than the plight of the slaves, so instead of abolition the law should be strengthened to protect the slaves and remove the abuses of slavery and suffering of the slaves (e.g. John Chrysostom, *Homilies on First Corinthians*, 40.5), while a few others called for the institution to be abolished (e.g. Gregory of Nyssa, *Fourth Homily on Ecclesiastes*). Abolishing slavery was not something that could be done overnight, as it represented a complete change of the economic system – and the wealthy slave owners who stood to lose the most were the ones with the largest clout in the Imperial court – but it was Christians in the late Roman period who led the charge to change the system. There are many examples of wealthy Christians releasing their slaves, or buying slaves so that they could set them free. However, the barbarian invasions of Western Europe, and constant Bulgarian and Islamic pressure in the East, set the cause back. As the church regained prominence in Northern Europe and political stability returned, the question of slavery again rose, and gradually over the course of the middle ages the Catholic abolitionist movement grew in strength and influence. By the high middle ages the practice of chattel slavery was outlawed by the Church across Northern Europe; for the first time in human history a society without slavery was established and endured, and it was accomplished by the medieval Catholic Church. In Southern Europe, slavery diminished, but remained, in part due to contact with the Islamic world where slavery was rampant. Muslim pirates raided Christian coasts as far away as Iceland in order to find slaves to sell at the market (not to mention the appalling trans-Saharan slave trade which brought far more death and misery to Africans than the trans-Atlantic trade); some Europeans responded in kind (though to a much lesser extent), trying to argue that the prohibition against slavery only applied to Christians. That small trickle turned into a flood as the explorers came forth from the newly liberated Iberian Peninsula, and a diabolical decree from one of the Renaissance bishops of Rome – in my view the worst bull ever issued by that church in that period, and it has some competition allowed them to enslave the native populations of the lands they discovered. Subsequent popes quickly reversed the ruling, and aside from that one lapse (when the Church was at its

Most, but not all. For example, it is difficult to read the accounts throughout Joshua or the commands given to King Saul[63] without feeling some pity for the Canaanites, the Hittites, the Amorites, the Perizzites, the Hivites, the Jebusites, the Amalekites, the Ammonites[64], the boys whom Elisha cursed[65] or the people at the time of Noah[66].

We must ask why we feel pity for these people: because we instinctively do not believe that they deserve death. They are no worse than us. Yet, this is not what the Bible records. With regards to the conquest under Moses and Joshua, these people were hopelessly immoral, and it was on account of that immorality that they were rightly judged[67]. There was no possibility of those societies reforming and changing; the hearts of those men had become no better than the hearts of beasts; and their existence would be a permanent temptation for the people entrusted with God's message (the were not completely eliminated, and the remnant that survived Joshua did constantly lead the Israelites astray: the Old Testament story is one of a people constantly abandoning God and His law to imitate their stronger neighbours, and then returning to Him. It is easy to see that without Joshua's campaign this temptation would have been stronger and perhaps permanently fatal, and without Israel and Judah there would have been no Christ).

Yet I agree that these people, bad as they were, were no worse than us. The Christian narrative is that we all deserve death; to be slain as the Amalekites were slain. We only do not all receive this death because of God's grace and mercy, to give us a chance to turn from our ways and repent. Before God, we do not have the right to life, for such a right could only come from merit, and we have no merit.

> There were some present at that very time who told him [Jesus] about the Galileans whose blood Pilate had mingled with their sacrifices. And he answered them, 'Do you think that these Galileans were worse sinners than all the other Galileans, because they suffered in this way? No, I tell

---

most corrupt and least Christian) the Roman Church has led a constant and vigorous opposition to slavery. Once again, many devout Roman Catholics begun working to end the practice of slavery (against those, largely arguing from secular sources such as Aristotle, that it should be maintained), and after about a century both the Church and King of Spain issued decrees outlawing the buying and selling of slaves. The American slave owners refused to surrender; Catholic Europe could no longer supply them, and that meant that there was a lucrative gap in the market which the Protestant nations of Northern Europe quickly took advantage of. The slave trade continued and grew stronger, until once again Christian opposition, this time led by Anglicans and Quakers, brought it to an end. In both the earlier Spanish and later British and Dutch slave trades, it was secularised merchants who pushed for the trade and the money to be earned from it, and devout Christians who opposed it. Not all Christians have been opposed to slavery – there have been times and places when most were in favour of the institution – but since the birth of the faith there has been a constant anti-slavery element in it (sometimes the majority, often the minority, but never zero), and for about 1700 years until the Christians decisively won the debate and the secular philosophers started to imitate them, almost all opposition to slavery was based on Christianity (the few Buddhists I mentioned who independently took up the cause being the only exception I am aware of). Thus the Church's historical relationship to slavery is nothing for Christians to be ashamed of. Those who criticise the Church generally look to eighteenth and nineteenth century examples (of which, regrettably, there were many to choose from), and ignore the wider sweep of Church history: that period was the anomaly. We should also not neglect that there are various different forms of slavery, some far worse than others; indeed some were regarded as beneficial for the slave, including by the slaves themselves. For example indentured or debt slavery, where someone could voluntarily sell himself and work until he had paid off his debts, under legal protection to avoid abuse, was the best recourse to the problem of debt known in the ancient world - one otherwise had to rely on handouts from one's neighbours, which was not sustainable over the long term, or would be left starving on the streets. Penal slavery was seen as a just punishment for some crimes, as an alternative to the likes of fines or prison sentences.

[63] I Samuel 15:2-3.

[64] 2 Samuel 8:2.

[65] II Kings 2:24.

[66] Genesis 6-7. Many might dispute the existence of Noah or even the other groups mentioned. I would simply say that we are considering the character of God as presented in the Bible; and these acts are part of that presentation. The question of whether such events occurred in practice is different.

[67] Genesis 15:16.

> you; but unless you repent, you will all likewise perish. Or those eighteen on whom the tower in Siloam fell and killed them: do you think that they were worse offenders than all the others who lived in Jerusalem? No, I tell you; but unless you repent, you will all likewise perish[68].'

By 'perish', of course, Jesus does not just mean their horrifying manner of death, but the suffering they subsequently faced in hell. Repentance need not save us from towers falling on us; but it can save us from hell.

God is just, and justice means giving each as they deserve. We should not complain when those who deserve death are given death; rather we should rejoice when they and we who deserve death are granted life. Some may consider this position distasteful; but the question is not what is distasteful to us, but what is true.

Now, this is not to say that we today have the right to commit such acts; on the contrary we will be condemned for them just as the Babylonians, unconsciously agents of God's justice[69] were condemned for the evil they committed when carrying it out[70]. It is for God to pronounce the sentence; we cannot, for we are just as deserving of death as those we condemn[71]. Furthermore, the argument above asserted that those condemned were irredeemably immoral, with no chance of repentance. That is no longer the case; since the coming of Christ, repentance and the chance of being transformed from immorality to morality is openly offered to everyone. We cannot say that anyone is hopelessly immoral, since we do not know how God is yet to act in their lives.

So if we feel horror at the destruction of the Amalekites then, unless that horror is simply the realisation that we or others may face the same fate, it is a fault in ourselves rather than in God. We do not have a right conception of human nature, of sin, or the righteous response to it. ○

Objection:

(1) There is a symmetry between goodness and evil; they are just opposite extremes on the same scale.
(2) Therefore in any argument, 'good' is interchangeable with 'evil'.
(3) Every argument that can be applied to support the existence of a good God can also be used to support the existence of an evil God.
(4) Christians reject the arguments that would support an evil God.
(5) Christians therefore also ought to reject the arguments that support a good God.

•

Response:

It should be obvious to the readers by now why this argument fails; but since I have frequently seen forms of it recently, it is worth refuting directly. There are two main flaws with the argument. Firstly, it regards God as being the ultimate embodiment of goodness. This is a bit over-simplified; but close enough that I can accept it for the sake of this response. Can there not exist a being which is the ultimate embodiment of evil? No; because existence is in itself a good. Any being which is evil in every respect cannot exist. Therefore the ultimate evil, the evil God, cannot exist. This, of course, assumes that existence is a good, and that existence is a property, but both of these were demonstrated above. Secondly, we saw above that God may be defined as the necessary being whose essence is his existence (and

[68] Luke 13:1-5.
[69] Jeremiah 25:9.
[70] Jeremiah 25:12.
[71] Matthew 7:1-2; Romans 2:1-4.

thus natural purpose is that He exists, along with everything that logically accompanies that); the necessary being is necessarily unchangeable in essence: pure actuality, without any potentiality; and the necessary being therefore has to be essentially good. This immediately rules out an evil God. Thirdly, the first premise of the argument is false. Evil and good are not opposite positions of what is essentially the same thing. There are many ways to be evil; but only one way to be good. Evil is seen as a deficiency in good; we cannot reverse this and say that good is a deficiency of evil. For example, the virtue of courage may be regarded as a deficiency in cowardice; but then, so is recklessness even more so. If courage is good because it is an absence of cowardice, why is recklessness not even better? Equally, is the most perfectly evil being a coward or reckless? Surely neither one should be regarded as being worse than another. Extending this for all the virtues, we conclude that there is therefore not one ultimate evil but an infinite number of equally ultimate evils; yet there is only one ultimate good. Therefore, there is no symmetry between good and evil, and the first premise of the objection is invalid. ○

CHAPTER 15

From Theism to Physics

15.1. Introduction

In chapter 11, I showed that, while classical physics (i.e. Newton's and the theory of General Relativity) were inspired by metaphysical positions contrary to classical theism, quantum field theory is consistent with the premises that underlie the classical definitions and conceptions of God; namely the concepts of actuality and potentiality; of form and matter; and the four causes, efficient, formal, material and final. I subsequently outlined that using deduction based on those premises, a conception of God consistent with Christianity (as well as Judaism and to a lesser extent Islam) necessarily follows. I described how the classical view of God could emerge from metaphysical premises which are consistent with modern physics.

To prove the equivalence of two ideas, A and B, it is necessary to show both that A implies B and that B also implies A. I earlier started from physics, and demonstrated that the existence of the God of classical theism is consistent with it. To complete my argument, I must also start from the premise that the God of classical theism exists, and show how modern physical theory might be deduced from that premise. I shall offer some initial thoughts towards that goal in this chapter. I will not be able to prove that the universe must be as we observe it; at various points I will need to make choices in my derivation, choices which God could but need not have taken. How far will I be able to go? That we shall see.

Metaphysics comes before physics. That means that we ought to be able to start from a metaphysical position, and use premises drawn from that position to reason towards a conclusion, namely a theory of physics. Since physical theory is expressed mathematically, this means that we will have to first of all translate the metaphysical theory into some abstract notation, which we can subsequently build on. Finally we should test the theory against experiment. My goal in this chapter is to begin with premises drawn from classical theism, and show that given this understanding of God we can expect the universe to be governed by a quantum field theory subject to the symmetries of the standard model, locality and renormalisability. My presentation in this chapter is therefore going to be the most mathematical of this book. I make no apologies for that: when the goal is a mathematical model, it is inevitable that the mathematics will be dense. I offer a short non-mathematical summary of the argument in section 15.10.

15.2. Premises

I begin with nine premises, and all but the last are either self-evident or arise from the classical theory of God.

(1) The God of classical theism exists as described in chapter 13;
(2) God has free will, only bound by the need to be self-consistent and consistent with God's nature (conjecture 13.3.5);
(3) God has an intellect, and is able to grasp abstract concepts and subsequently actualise them (conjecture 13.3.12);
(4) The opening statement of the Bible is accurate:

> In the beginning, God created the heavens and the earth. The earth was without form and void, and darkness was over the face of the deep[1];

i.e. that there is a material world apart from God, of which God is the ultimate cause (definition 13.1.12).

(5) Whatever God can create, He can also destroy;

(6) The universe is changeable, and it can exist in a large number (possibly an infinite number) of possible states;

(7) There are various different kinds or types of matter (corresponding to the fundamental particles – at this point in our discussion there might be a finite or infinite number of these types of particles, so all I assume here is that the set of all irreducible (non-compound) material objects can be divided into various subsets, and that the number of these subsets is greater than zero);

(8) We can compare two states of the universe and see if they are the same or different;

(9) Space and time exist and can be mapped onto a $3+1$ dimensional Riemann geometry, and that the state of matter can be represented at least in part symbolically and probabilistically (these lead to the idea that physics can be represented at least in part mathematically). Because of the rational nature of God, the mathematical and geometric representations representing reality should be continuous and free from infinities. Just as reason is a chain of small steps from premise to conclusion without any sudden jumps from one step to another, so should the evolution of the universe proceed in small steps from an initial state to a final state without any sudden jumps in its state or our abstract description of that state. Our knowledge of the universe at one moment should be informed by our knowledge of the universe at the previous moment, with the similarity being stronger the smaller those moments are. Without this, it would be impossible to use reason about the universe; but God is an intellect, and thus considers the universe using reason.

In premise 1, I include in particular the concepts that God is outside space and time (and any other similar quantities), and therefore relates to each material object in the same way regardless of its location or time, omnipotent, omniscient, and also that God is unchanging and not only the creator but also the sustainer of everything else.

Premise 5 includes the idea that God is capable of creating, sustaining and destroying matter (either directly or indirectly through an agent). Premise 6 follows from the idea that only God is a being of pure actuality; since everything else contains potentia, it must be possible to for it to undergo change. Since matter exists in various potentia, we know that not all possible material beings are identical, although the idea that there are different types of matter goes beyond this.

Premise 9 is perhaps the hardest to justify in terms of classical theism, although it is not contrary to it. One key idea behind the wording of this premise is that we want the formalism we use to understand matter to be as general as possible. If we choose something specific, then we might get it right or we might get it wrong – but this is not something we can decide when laying out our premises; using a general formalism, however, will allow us to later focus on any one of the specific formalisms at the end of the argument. So our plan would be to adopt the most general abstraction consistent with the rest of the premises, and take specific cases at the last possible moment in the reasoning. The most significant place where I broke this rule is where I specified three extended spatial dimensions, which is sufficiently evident from observation that I think that I can safely add it to my axioms.

We have stated that matter is capable of change; there must therefore be some parameter which measures that change, so when this parameter is at one value the system is in one state, and another value the system might be in another state. This parameter counts moments in the sequence of causality and therefore can be represented by a numerical scale. The next question is whether this scale is continuous or discrete. However, if the universe

[1]Genesis 1:1-2.

is capable of being understood rationally, the state of the universe at one instant must be similar to the state at the previous instant along this scale. If the change is to be smooth and continuous (the alternative being that it changes in discrete jumps, which spoils the idea that the changes can be understood rationally since we would not then be able to say for certain that the change in X was the cause of the change in Y), then this measure of change must be a continuous scale since only then is the concept of continuity defined, which means that time can be represented by a line and thus geometrically. Equally, whatever it is that is changed must be able to be measured by some continuous quantity, which gives us the idea of a d dimensional space. This space might be either compact (e.g. rolled up into a circle), or non-compact (e.g spread out over a line), or a combination of the two. Either are, I think, permissible, though in practice we will need both three non-compact dimensions, and a number of compact dimensions. The definition of time as a measure of change restricts it to have a linear topology. In either case, we can again represent these quantities geometrically. Because of these contrasting definitions – one as the measure of change itself, and the other the measure of what is changed – space and time must be distinguished from each other, and there are several options, including the Galilean geometry of $\mathbb{E}^3 \otimes \mathbb{E}^1$ and the hyperbolic (Minkowski) geometry $SO(3,1)$. However, the hyperbolic geometry is more general, as it reduces to the Galilean geometry in a certain limit (to move ahead of myself, the $c \to \infty$ limit, where c is the speed of light), so if we consider hyperbolic geometry we can always later revert to the special case of Galilean geometry if required, but if we use Galilean geometry as the basis of our understanding of space and time we lack the freedom to switch to the hyperbolic geometry. This geometry is only to determine the signature and shape of the metric tensor. Once we convert the geometry into a coordinate system, we can always consider non-inertial coordinate transformations, which forces us to generalise to a Riemann geometry with a hyperbolic signature of the metric; once again, this is the most general geometry consistent with the assumed topology and metric signature, so if in practice we need something else we can always move from the most general solution to something more specific. Of course, there is no immediately obvious reason why we should restrict ourselves to three dimensions, or a finite speed of light – those are (at this point of the argument) arbitrary choices which God would have to make when designing the universe (although they are in practice evident from observation).

The other part of premise 9 is that the universe can be understood abstractly. This follows from God's omniscience; if God understands everything about the universe, then it is in principle possible to come to an intellectual understanding of the universe, which means that there must be some abstract layer to the universe which is actualised by some concrete principle (matter). The geometrical nature of space and time implies that this understanding must be at least in part geometrical, and the rational nature of God implies that there must be a logical progression from premise (the state of the universe at one moment) to conclusion (the state at the next moment), implying that spatial and temporal changes in the universe can be broken down into a set of smooth and continuous changes. God's rationality implies that the universe can be understood logically, but since probability is the generalisation of logic, we will use a formalism based around probability. This might allow for a deterministic physics (by forcing those probabilities to be either one or zero), but if we started assuming that physics was deterministic and constructing the formalism accordingly, we would not be able to later switch to an indeterminate physics.

In addition to these premises, I will make one more in this chapter, which is that God is impartial to mankind, in the sense that God's actions are not influenced by the state of mankind, so the rule in premise 1 that God's actions affect each particle in the same circumstances in the same way regardless of its location or time is not broken by any desire to help or harm mankind or individual people. The assumption is that wider concerns on earth do not alter the circumstances surrounding God's decisions concerning individual electrons and quarks. God can, of course, still feel pain (or whatever is analogous to feeling

pain in the mind of God) at mankind's evil, but this assumption prevents Him from acting in response to that pain. My use of the word *impartial* here thus might be misleading, but I have not been able to think of an alternative that is closer to describing the concept.

This assumption is contrary to the conclusion of the previous chapter (conjecture 14.7.1). I do not believe it to be correct. But I will accept it for now to provide us with a first approximation to nature. It is common in physics, when faced with a seemingly insoluble problem, to start by finding a similar but soluble problem, solve that, and then systematically correct our answer so we approach the true solution of the actual problem we were given by nature. What I intend to do is no different from that. We know that God has to balance respecting the freedom of His creation and acting to prevent and minimise other forms of evil. The first side of the scale demands that God will seem to act impartially to mankind, the second side demands that God is constantly active, and the optimal situation will be a balance between the two. I am here assuming that that balance tends towards the 'respect freedom' side of the scale, so I will first consider the universe as though God were impartial, and later on add in the effects of His compassion as a small perturbation.

15.3. Quantum Field Theory

From these premises, we can see several things.

(1) The universe is apart from God and distinct from him. There are therefore objects which can exist which are not God; in particular material objects, the union between form and matter. We can tell that there is form, because beings within the universe are planned out in God's mind before their creation, so each being must be describable by an abstract concept; and there must be matter because once actualised the being exists apart from God, so therefore the form must be capable of being instituted in something other than an intellect, which we denote as matter. Since the objects are capable of change (premise 6), this form must contain a description of a set of possible states, since God is able to understand the being both before and after the change.

(2) There is a state of the universe which is formless and void (see premise 4), which I will denote as $|0\rangle$.

(3) There are at least one non-formless and void states which the universe could be in. These we may denote as $|\psi_n\rangle$, where each different index n represents a different possible state of the universe. For example, $|\psi_1\rangle$ might denote a universe where there is just one particular electron in it (we need not assume that electrons exist, any type of matter will do), $|\psi_2\rangle$ a state with two distinct electrons in it, and so on. One of those states, say $|\psi_{H,E}\rangle$ will represent the heavens and the earth as they were at the moment of creation.

(4) There is some operation by which we can describe God's construction of the heavens and the earth (and any other state) from the formless and void state, i.e. there is some operator $\hat{\psi}^\dagger_{H,E}$ which acts on $|0\rangle$ to give $|\psi_{H,E}\rangle$. In symbolic notation, this reads, $|\psi_{H,E}\rangle = \hat{\psi}^\dagger_{H,E}|0\rangle$. The same is true for any other possible state of the universe. This operator must exist; since God did transfer the universe from a formless and void state to a heaven and earth state, this process must be possible. If the process is possible, then it is helpful to have some symbolic notation to represent it. Using our notation for the individual states, we need some object which acts on the void state and emits the heaven and earth state; and this operator describes the whole process of creation. $|\psi_{H,E}\rangle = \hat{\psi}^\dagger_{H,E}|0\rangle$ states that 'formless and void' can turn into 'heavens and earth' in certain conditions, and it describes this complete process symbolically. It doesn't describe the microscopic details of how that is accomplished; nor does the notation even mention God, although obviously God is there in the background motivating this whole procedure, and we can refine the properties of these creation operators using the known attributes of God.

(5) There is also some operator which I will denote as $\hat{\psi}_{H,E}$ which reverts from the heaven and earth state to the formless and void state (premise 5), so $\hat{\psi}_{H,E}|\psi_{H,E}\rangle = |0\rangle$, and the same is also true for any other possible state of the universe. It is therefore possible for God to change any one state of the universe into any other, if nothing else by combining the operators $\hat{\psi}_2^\dagger\hat{\psi}_1$ acting on the initial state $|\psi_1\rangle$.

CONJECTURE 15.3.1. *It is possible to use an abstract notation to describe the various possible states of nature.*

Proof:

(1) God has an intellect.
(2) The intellect is something which is capable of grasping abstract concepts.
(3) God understands the universe, in each possible configuration, in His intellect 'prior' to creating it.
(4) It must be possible to represent the universe in abstract thought.
(5) Everything which is capable of grasping in abstract thought can be grasped, at least partially, in some abstract notation.
(6) Each possible configuration of matter can be represented, at least partially, in an abstract notation.

■

Discussion:
The word 'prior' should be taken as meaning earlier in the sequence of efficient causality rather than earlier in time.

Note how the idea that the universe is understandable follows from God's intellect and omniscience; since God understands it, it must be possible that it can be understood. Obviously, we are short of saying that we can understand it in full, since our intellect is more limited than God's. However, even though we have a more limited intellect, we can still expect to reach a partial understanding; for if we could not understand anything we would not have an intellect. Now a partial understanding might involve only being able to grasp a few specifics but not general concepts, only being able to grasp general concepts but not specifics, or some combination of these. We should not assume *a priori* until proven otherwise that our intellect is limited in either the respect of being able to grasp specific details or general concepts; we leave the possibility that it can understand either or both types of knowledge open.

Having stated that we can describe the universe in some notation, we have some freedom in how to do so. We need some symbol to represent each possible state of matter; obviously quite how we do this is arbitrary, but it needs to be extensible enough to at least in principle be able to cope with each of the infinite number of situations. Secondly, we need our notation to be able to cope with the concepts of change and comparison; this follows from our premises that both of these are possible. □

Comment:
I have written these state vectors in a discrete notation, with each label as an integer. I have done this because it is easier to visualise and write down. In practice, we would want to use a continuous notation, since space is continuous. This is easy to do: instead of discrete indices using integers to distinguish between one state and another, we can use a set of real numbers, replacing the sums with integrals. We can also use both real numbers and integers to label the states. I am just using integer labels in this introduction because it is simpler to explain.

Also note that to start with, I will use a temporal picture of physics, where we consider each moment of time separately, rather than a four dimensional picture. □

Now each of these states represents a set of information; $|\psi_1\rangle$ not only says that there is a particular lump of matter present in the universe, it describes the absence of every other possible lump of matter. In particular, since we can describe each of the properties of an electron with either a discrete or continuous set of numbers, each of the $|\psi_n\rangle$ states can be written as a series of numbers (for example, we might use four numbers to represent its momentum, one for its spin, and another one to describe whether or not it exists); in short, the mathematical object known as a vector (defined over some particular space; not the usual continuous Hilbert space we are used to, but something similar). Now between them all the $|\psi_n\rangle$ states and $|0\rangle$ by definition describe every possible state of the universe, so they form a complete set of basis vectors.

The other element we need is that God has an omnipotent free will. God can change the universe from any state to any other state, and this process is unpredictable as a consequence of God's will being free, the process being restricted only by the needs of self-consistency and God's nature. This means that the evolution of the universe must be (to us) probabilistic rather than deterministic. Given the universe in one state, God can change it to any other logically possible state in the next instant in time.

Objection:
The idea that God is free to move any state of the universe at one time to any other at the next moment implies that there is not necessarily any continuity in physics – anything can turn into anything else. This is consistent with a free-for-all causality, perhaps similar to that proposed by David Hume (more technically occasionalism[2]), but not the more structured Aristotelian causality which was used earlier to describe the nature of God. Since we are here assuming that same nature of God, this implies a contradiction. •

Response:
The statement was not that God is free to do anything, but God is free to do anything that is self-consistent and consistent with His nature. This might naively seem to be a trivial distinction, but, as we shall see as we work through the details, it is in practice very restrictive. These restrictions will manifest themselves in certain symmetries of nature and locality principles, which in turn will imply various conservation laws, which lead to some vision of causality similar to that of Aristotle and contrary to Hume and away from occasionalism. ○

So suppose that at one moment in time, t_1, the universe is in a state $|\psi(t_1)\rangle = |\psi_A\rangle$. We may ask the question what state will the universe be in in some later moment of time, t_2. The answer is that we don't, and can't know. God is free to move the universe into whatever state He feels like, baring the usual constraints. However, what we can then say is that there will be some likelihood that the universe will remain in state $|\psi_A\rangle$, and some likelihood that it will move to $|\psi_B\rangle$ and some likelihood that it will move to $|\psi_C\rangle$ and so on (since everything that isn't a certainty can be expressed as either a probability or likelihood or some other measure of uncertainty). We therefore need some means to introduce some indeterminate language into our formulation. It will do this by writing $|\psi(t_2)\rangle = \sum_n c_n(t_2)\,|\psi_n\rangle + c_0(t_2)\,|0\rangle$, where the coefficients c_n will in some respect (we will discuss how later) represent the likelihood of the universe being in the associated state. Now, one important question is whether these c_n are real or complex numbers; for the moment, to keep things general, I will allow them to be complex, and return to this question later.

[2]The belief that material objects cannot be secondary causes; that everything is caused by God alone

CONJECTURE 15.3.2. *We can introduce some comparison operator $\langle\psi_n|$ which acts on the state vectors $|\psi\rangle$, is associative, and gives the coefficient c_n if $|\psi\rangle = \sum_n c_n |\psi_n\rangle$.*

Proof:
We assumed as an axiom (premise 8) that we can compare different states of the universe, that is to say there is some operator which can identify whether or not we are in a state $|\psi_n\rangle$. We denote this process with the state vector $|\psi_n\rangle$, and we choose to define this operator in such a way that $\langle\psi_n|\psi_n\rangle = 1$ and $\langle\psi_n|\psi_m\rangle = 0$ for $n \neq m$. 1 here means that they are in the same state, and 0 that they are not. This provides the definition of the $\langle\psi_n|$ notation.

For a general vector $|\psi\rangle$, we are uncertain which state it is in, though we know that it must be in some state. Suppose that we know that the state of the universe represented by $|\psi\rangle$ is in either state $|\psi_1\rangle$ or state $|\psi_2\rangle$. Thus $\langle\psi_n|\psi\rangle = 0$ for n anything except 1 or 2. If n is 1 or 2, then $\langle\psi_n|\psi\rangle$ cannot be 0 because $|\psi\rangle$ might be in state $|\psi_n\rangle$, and it cannot be 1 because $|\psi\rangle$ might not be in that state. Instead, we should write $\langle\psi_n|\psi\rangle$ in some way that parameterizes our uncertainty of the state of $|\psi\rangle$, and we can give this the value c_n, so $\langle\psi_n|\psi\rangle = c_n$. We still have a great deal of freedom in constructing our notation, so we can define $\langle\psi_n|$ to be associative, so that $\langle\psi_n|(c_1|\psi_1\rangle + c_2|\psi_2\rangle + c_3|\psi_3\rangle + \ldots) = c_1\langle\psi_n|\psi_1\rangle + c_2\langle\psi_n|\psi_2\rangle + c_3\langle\psi_n|\psi_3\rangle + \ldots$. This construction makes the working considerably simpler. Given that, by definition, $\langle\psi_n|\psi_m\rangle = \delta_{mn}$ where $\delta_{mn} = 1$ when $m = n$ and 0 otherwise, this means that $\langle\psi_n|\psi\rangle = c_n$. Thus the c_n coefficients represent a parameterization of our uncertainty about the future evolution of the universe. ■

CONJECTURE 15.3.3. *We can also introduce a state $\langle\psi|$ which can be used to compare whether or not a state $|\psi'\rangle$ is in state $|\psi\rangle$.*

Proof:
This is a little more complex, because $|\psi\rangle$ is not a definite state, but only represents the case where we are uncertain which state it is in. Let us introduce a second of these states $|\psi'\rangle$. With regards to the comparison between $|\psi\rangle$ and $|\psi'\rangle$, we can say that we know one of three things; a) that they are definitely not in the same state (this might be the case if $|\psi\rangle$ is known to be in an even numbered basis state, and $|\psi'\rangle$ is known to be in an odd-numbered basis state; b) they are definitely known to be in the same state (for example, if we want to compare $|\psi\rangle$ with itself); and c) we are unsure whether or not they are in the same state. Our notation needs to be able to handle all three of these cases. We can thus introduce the comparison state $\langle\psi|$, which is defined according to the rules:

- $\langle\psi|\psi'\rangle = 0$ if $|\psi\rangle$ and $|\psi'\rangle$ are known to be different.
- $|\langle\psi|\psi'\rangle| = 1$ if $|\psi\rangle$ and $|\psi'\rangle$ are known to be the same.
- $|\langle\psi|\psi'\rangle| \neq 0, 1$ if we are unsure whether or not $|\psi\rangle$ and $|\psi'\rangle$ are different or the same. The value of $\langle\psi|\psi'\rangle$ then in some sense represents the likelihood that these two states are the same.

Now, just as we wrote $|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle + c_3|\psi_3\rangle + \ldots$ as a means of expressing our uncertainty of which state it was in, so we should express $\langle\psi|$ in a similar way. $\langle\psi| = \bar{c}_1\langle\psi_1| + \bar{c}_2\langle\psi_2| + \bar{c}_3\langle\psi_3| + \ldots$. This is because we can apply $\langle\psi|$ to a basis vector $|\psi_1\rangle$. Now, if there is a certain likelihood that $|\psi\rangle$ is in the same state as $|\psi_1\rangle$, which we denote as $\langle\psi_1|\psi\rangle$, then there must also be a certain likelihood that $|\psi_1\rangle$ is in the state $|\psi\rangle$, which we denote as $\langle\psi|\psi_1\rangle$. These two likelihoods are obviously related to each other, but in what way is not yet clear (we have not yet, after all, precisely defined what we mean by 'likelihood'). However, if we know that $|\psi\rangle$ is definitely not in the state $|\psi_4\rangle$, then both $\langle\psi|\psi_4\rangle = \langle\psi_4|\psi\rangle = 0$. This makes the notation

$$\langle\psi| = \bar{c}_1\langle\psi_1| + \bar{c}_2\langle\psi_2| + \bar{c}_3\langle\psi_3| + \ldots \tag{635.1}$$

a natural one to use. $\bar{c}_1$ will be related to c_1 in some way, but we haven't yet determined what this relationship is. ■

CONJECTURE 15.3.4. *The coefficients c_1 and $\bar{c}_1$ cannot represent probabilities.*

Proof:
In mathematics, the probability is defined according to certain axioms. The probability of A, $P(A)$ is a real number between 0 and 1 which describes our knowledge of whether the proposition A is true, with 1 indicating that it is true, and 0 that we are certain that it isn't. The language of probability is natural for this physical picture we are building up, because we have various statements we can investigate – is $|\psi\rangle$ in the same state as $|\psi'\rangle$? – which we have uncertain knowledge about, and we want to classify that uncertainty. However, the mathematical definition of probability also requires a second axiom, namely that if both A and B cannot be true, then the probability of being in either state A or state B is the sum of the probability that it is in state A plus the probability that it is in state B. Symbolically, this is written as: if $P(A \cap B) = 0$ then $P(A \cup B) = P(A) + P(B)$. Now let us see if our coefficients c satisfy this requirement.

Suppose we have a state, $|\psi\rangle = c_A |\psi_A\rangle + c_B |\psi_B\rangle$ (I'll use A and B to denote the fundamental orthogonal basis states in this example, to keep the probability notation more convenient). Now we also know that $\langle\psi|\psi\rangle = 1$, which implies that

$$\bar{c}_A c_A + \bar{c}_B c_B = 1. \tag{636.1}$$

Let us write $P(A)$ as the probability that $|\psi\rangle$ is in state $|\psi_A\rangle$ and $P(B)$ as the probability that $|\psi\rangle$ is in state $|\psi_B\rangle$. By construction, ψ cannot be in both states $|\psi_A\rangle$ and $|\psi_B\rangle$ simultaneously, so $P(A \cap B) = 0$. Now, we also know (in this example) that ψ must be in either state A or B, so we require that $P(A \cup B) = 1$. If c_A were the probability $P(A)$, then this would mean that $c_A + c_B = 1$, which contradicts equation (634.1) unless $\bar{c}_A = \bar{c}_B = 1$. But this makes no sense, because we know that $\langle\psi|\psi_A\rangle \neq 1$, which is an expression that we are not certain that ψ is in state ψ_A. This means that $P(A) \neq c_A$.

Of course, this example just used a two state system. However, it is straight-forward to extend it to any number of states. These multi-state systems can be considered in a similar way; the algebra is a little more complicated, but the basic conclusions still hold. ■

Discussion:
However, equation (634.1) does suggest the solution. If we write $P(A) = \bar{c}_A c_A$ and $P(B) = \bar{c}_B c_B$, then the addition rule is satisfied. Indeed, given (634.1), it is difficult to think of any other way we can construct a probability from c_A and $\bar{c}_A$ that satisfies the addition rule. Of course, we also require that $P(A)$ is real and positive, and this requirement is satisfied if $\bar{c}_A$ is the complex conjugate of c_A, i.e. (expressing the same thing in symbolic notation, where † indicates taking the complex conjugate) $\bar{c}_A \propto c_A^\dagger$ (this applies whether or not c_A is restricted to be a real number; as I said, we will keep things as general as possible for now), where the constant of proportionality is real and positive. Since both $|\langle\psi|\psi_A\rangle|$ and $|\langle\psi_A|\psi\rangle|$ represent likelihoods which have the same maximum value of 1, the constant of proportionality must be 1 and $\bar{c}_A = c_A^\dagger$. Now the remaining requirements for a probability fall into place. We said that $c_A = 0$ represents certainty that $|\psi\rangle$ isn't in state $|\psi_A\rangle$, which implies that $c_A^\dagger c_A = P(A) = 0$, so the probability works out as it should. Similarly certainty that it is in the required state, $P(A) = 1$, implies that $c_A^\dagger c_A = 1$, which is consistent with our earlier notation. Thus if we are to express our conclusions in terms of probabilities, we are left with no choice but to accept the following conjectures. □

CONJECTURE 15.3.5. *Given a state vector* $|\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle + \ldots$ *and a comparison vector* $\langle\psi| = \bar{c}_1 \langle\psi_1| + \bar{c}_2 \langle\psi_2|$ *such that* $\langle\psi|\psi\rangle = 1$, *then* $\bar{c}_1 = c_1^\dagger$.

This means that for any general state vector $|\psi\rangle$, the corresponding $\langle\psi|$ is its Hermitian conjugate.

CONJECTURE 15.3.6. *Given a state vector* $|\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle + \ldots$, *the probability that the object described by* $|\psi\rangle$ *is in state* $|\psi_1\rangle$ *is* $c_1 c_1^\dagger$, *and consequently* $\sum_n c_n c_n^\dagger = 1$.

Discussion:
Note that the observable of physical importance is the probability, $c_1^\dagger c_1$. When we perform an experiment (or rather a lot of measurements of the same observable), what we measure is the frequency distribution for a particular outcome, which tends to the probability distribution in the limit of an infinite number of independent measurements. c_1 itself is not directly observable. Now, c_1 is a complex number, and if we rotate it by a complex phase, $c_1 \to c_1' = e^{i\theta} c_1$ for some arbitrary real number θ, then the probability distribution is unchanged, $c_1^\dagger c_1 = {c_1'}^\dagger c_1'$. This abstraction is intended to describe physics, so if each of these possible choices describes the same probability distribution, then they are each equivalent abstractions. The difference between c_1' and c_1 is just something we have introduced when creating the abstraction we use to represent physics. It is not necessarily (at this stage of the argument) a part of the real world. Therefore, our final theory should respect this as a symmetry in some way: the absolute phase of the states should disappear when we map back from the abstraction to the physical world (and our experiments). The relative phase between two states might, however, remain important.

Having discussed what these amplitudes c are not, I should take a moment to discuss what they are. Our goal in all this is to make predictions about the future. God is free to act how He likes, so we cannot predict with complete certainty what is going to occur. However, we can still say how likely each option is. Parameterization of uncertainty need not be expressed as a probability. These amplitudes thus represent parametrisations of uncertainty, conditional on an initial configuration, of what the quantum states are going to in a future circumstance. The quantum states themselves, for example $|\psi_1\rangle$ play a dual role. Firstly, they are a representation of possible physical states. Every physical state will be in a one to one mapping with one of these abstract states. But often we don't know which state the system is in, or which state it is going to be in after a particular interaction. The same notation and construction can be used to both describe actual reality and to assign degrees of confidence to make predictions about possible future events. Strictly speaking these are different functions, so when we write down a quantum state, we should be clear whether it is to be used to record the result of a measurement or to express our degree of certainty about a prediction of the results of a currently unknown measurement. The same notation is used for both.

So the quantum system is always in a given state. We don't usually know which state that is, which means that we have to express that uncertainty by writing down all the possibilities weighted by the appropriate likelihoods, and summing over them. That is then a representation of our knowledge of the current state of the particle. Then, when we want to predict the results of a particular measurement conditional upon what we currently know about the system, expressed as a quantum amplitude, we can extract that information from the representation of our state of knowledge.

Those who characterise interpretations of quantum physics generally distinguish between ψ-epistemic and ψ-ontic interpretations. In a ψ-epistemic interpretation, the wavefunction refers to our state of knowledge of the system. In a ψ-ontic interpretation, the wavefunction represents the actual physical parameters of the system. Both classes of theories have their problems. In general, a ψ-ontic theory struggles with the double slit experiment, and implies that the particles are in some respects non-local, having to go through both slits at once, contradicting what we observe and what the theory in other respects demands, namely that the particles are localised in one particular place. ψ-epistemic interpretations, on the other hand, are generally thought to be ruled out because they give incorrect predictions for some experiments involving entanglement. My interpretation doesn't really fall into either category, or rather it combines elements of both. This can be understood in terms of the distinction between substance and event causality. When used to represent the state of substances, the quantum state is treated as a representation of an actual physical state. When used to make predictions about future events, the quantum state is treated as epistemic, containing information concerning several possibilities, only one of which would occur in practice. □

Thus, given our interpretation that this product in some way represents a probability that the two sides are in the same state, we have shown that the product must satisfy the properties of linearity, positivity, and commutativity. Mathematically, this means that the objects $|\psi\rangle$ lie in some inner product space. We can therefore bring into play all the known mathematical machinery related to such spaces. In particular, we can write any operators in this space in the form $H = h_{nm}\,|\psi_n\rangle\langle\psi_m| + h_{n0}\,|\psi_n\rangle\langle 0| + h_{0m}\,|0\rangle\langle\psi_m| + h_{00}\,|0\rangle\langle 0|$. For example, the creation operator for heaven and earth is given by $\psi_{H,E} = |\psi_{HE}\rangle\langle 0|$, while the destruction operator for heaven and earth is given by $\psi_{H,E} = |0\rangle\langle\psi_{HE}|$.

CONJECTURE 15.3.7. *The possible set of basis vectors is not unique.*

Proof:
A basis vector is defined as one of a complete set of irreducible orthonormal objects which can between them be used to represent everything in the physical world. A basis is the set of a consistent group of basis vectors. By saying that the basis is not unique means that there are numerous equivalent ways in which we can represent physical reality.

The basis vectors we have used to define the various individual states of the universe are defined according to two properties. Firstly, completeness, which states that every possible state of the universe is described by one state or another, and secondly orthonormality, which states that $\langle\psi_n|\psi_m\rangle = \delta_{nm}$. Given that we are permitting linear combinations of these states to describe our uncertainty, this leaves open the possibility that there might be numerous different basis states which satisfy these conditions. The choice of basis is a (necessary) artefact of our mapping from physical reality to an abstract representation of it; it is something we introduce which is part of the formalism but not part of the underlying reality. Therefore, as long as we consistently use the same basis throughout the calculation, the final result as mapped back to reality should be the same regardless of which basis we use.

So, given one set of basis vectors $|\psi_n\rangle$, we can consider a rotation to a different basis $|\psi'_j\rangle$, with

$$|\psi'_j\rangle = \sum_n R_{jn}\,|\psi_n\rangle\,, \tag{638.1}$$

for some matrix R_{jn}. I will generally omit the summation sign for repeated indices. Now, we need to check that the two conditions of orthonormality and completeness can be satisfied.

First of all, orthonormality,

$$\begin{aligned}\langle\psi'_k|\psi_j\rangle =& R^\dagger_{mk}R_{jn}\langle\psi_m|\psi_n\rangle \\ =& R^\dagger_{mk}R_{jm}.\end{aligned} \tag{639.1}$$

We take the Hermitian conjugate of the comparison vector because that's what we do when we convert from a state vector to a comparison vector. The orthonormality condition is satisfied if

$$\sum_m R_{jm}R^\dagger_{mk} = \delta_{jk}. \tag{639.2}$$

This is well known to be satisfied for a class of matrices known as unitary matrices. Thus if R is unitary, then the orthonormality condition is satisfied. Also note that every unitary matrix is invertible, which means that we can also rotate from the $|\psi'\rangle$ basis to the original $|\psi\rangle$ basis.

The second condition of completeness can be proved as follows. The first basis is complete, which means that we can write any possible state vector, $|\phi\rangle$ as a linear combination of those basis vectors.

$$|\phi\rangle = \sum_n c_n |\psi_n\rangle . \tag{639.3}$$

If we rotate to the new basis, we have,

$$|\phi\rangle = \sum_{nj} c_n R^\dagger_{nj} |\psi'_j\rangle , \tag{639.4}$$

which means that we can introduce a second set of coefficients $c'_j = \sum_n c_n R^\dagger_{nj}$. We then have,

$$|\phi\rangle = \sum_j c'_j |\psi'_j\rangle , \tag{639.5}$$

which means that $|\phi\rangle$ can be written as a linear combination of these new basis vectors. Since this applies for any choice of $|\phi\rangle$, every possible state vector can be written in the new basis, as long as the original basis is also complete and the new basis contains the same number of vectors as it.

Thus the $|\psi'\rangle$ basis is an equally good choice to describe our knowledge of reality as the original $|\psi\rangle$ basis.

■

Discussion:

It can be asked what these various different representations mean. It is easy to think of the state of the universe in terms of (for example) the locations of different particles, so $|\psi_1\rangle$ means that there is an electron in one particular place, $|\psi_2\rangle$ another electron at a different place, and so on. In this naive representation, there would be one particular basis state which corresponds to each possible set of particle positions, momentums, and any other properties of importance. However, because we cannot predict the future, we have to write our knowledge of the future as a linear representation of these physical states. So, for example, if $|\psi_1\rangle$ represents a universe which has only one electron in one location, and $|\psi_2\rangle$ an electron in a different location, then $\frac{1}{\sqrt{2}}(|\psi_1\rangle + |\psi_2\rangle)$ where there is a 50% probability that we will be in the first universe, and a fifty percent probability we will be in the second universe. This uncertainty is necessary because of the freedom of God's will, meaning that we cannot know the future with certainty. Thus the ability to rotate the basis is dependent on our necessary uncertainty of knowing how the universe will evolve. There might well in practice be one particular basis which corresponds precisely to there being an electron here, and another electron there, and so on, rather than a mixture of probabilities. However, since our knowledge of the future is necessarily uncertain (and our knowledge of the present

in practice uncertain), any state vector we use in practice will not be one of those idealised basis vectors, but the linear superposition of them. Thus even if there is some idealised basis corresponding to definite positions of particles, it is not particularly helpful nor necessary to use it for the sort of questions we want to ask. □

CONJECTURE 15.3.8. *If A is a state defined using basis R_1 and B a state defined using a basis R_2, the probabilities $P(A \cup B)$ and $P(A \cap B)$ only follow the standard rules for combining probabilities if the two bases R_1 and R_2 are identical.*

Proof:
For example, we might define the two bases as the eigenstates of operators H_1 and H_2, and they would be equivalent if the two operators had the same eigenstates (though possibly with different eigenvalues).

Premise 0 for probability, defined in section 2.2, defines probability in terms of a set of states which have to be irreducible and orthogonal. The eigenstates of one operator H_1 satisfy this property, but once we introduce the second operator, in general the eigenstates will no longer be orthogonal. We can only define a quantity $P(A \cup B)$ if A and B are members of the same initial set of basis states. This means that the probability can only be defined if A and B are either identical or orthogonal to each other (or, if they are a set of the irreducible basis states, their members are either identical to or orthogonal to each other). In general the eigenstate bases defined from different operators will not be orthogonal, which means that quantities such as $A \cup B$ violate premise 0 of the definition of the probability, and thus $P(A \cup B)$ is not a probability but undefined. Therefore the usual conclusions of probability theory, such as $P(A \cup B) \leq P(A) + P(B)$ need not be valid.

Before deriving probabilities from the probability amplitude, we must first select an eigenstate basis, and consistently use the same basis throughout the calculation. If we have an operator which represents an observable we want to measure whose eigenstates are not equivalent to the eigenstate basis chosen to define the probability, we must first rotate the operator eigenstates into linear suppositions of the probability eigenstates; only after doing this can we compute expectation values or probabilities of being in a particular state.

In quantum physics, the amplitudes obey the addition rule $|A \cup B\rangle = |A\rangle + |B\rangle$. This is not a probability because the amplitude is not positive definite. The square of the amplitude obeys an addition rule if the states $|A\rangle$ and $|B\rangle$ are orthogonal (i.e. $\langle A|B\rangle = 0$), and since this then obeys the addition, positivity and normalisation rules we can treat the combined amplitude using the rules of combined probability. We can still express the square of $|A \cup B\rangle$ as a probability (since we can always find some basis which has a irreducible basis vector proportional to $|A \cup B\rangle$), but it is not the combined probability of states A and B subject to the addition rule (the concept of a combined probability is meaningless unless A and B are orthogonal). We have to treat $|A \cup B\rangle$ as though it were an irreducible basis state. ■

CONJECTURE 15.3.9. *There exists a creation operator, $\hat{a}^\dagger$, which describes the process of God creating a new particle.*

Proof:
From His omnipotence and free will, God has the ability to create new particles as he sees fit (corollary to premises 1 and 5). Since the basis of states is complete, and covers all possible configurations of our knowledge of the material universe, there is one state which describes the universe just before that particle is created, and another which describes the

universe just after that particle is created. Let us write the first state as $\left|\psi_n^{(0)}\right\rangle$, the index n representing the numerous possible ways the universe can lack that particle but where it could in principle be created, and, $\left|\psi_n^{(a)}\right\rangle$ the corresponding state of the universe which contains that particle. One of the $\left|\psi^{(0)}\right\rangle$ states is the vacuum state $|0\rangle$. We are here operating in a suitable basis where these states are orthogonal to each other. In this case, the operator $\sum_n \left|\psi_n^{(a)}\right\rangle \left\langle\psi_n^{(0)}\right|$ is an operator which corresponds to the act of creating that particle. When applied to a given state vector $|\phi\rangle$, this replaces each state which lacks that particle which one that possess it. Thus this operator describes the act of God creating a particle, and we can label it as $\hat{a}^\dagger$. ■

CONJECTURE 15.3.10. *Creation operators either commute or anti-commute with each other.*

Proof:
If we take some initial state $|1\rangle$ and apply two (possibly) different creation operators $\hat{a}_n^\dagger$ and $\hat{a}_m^\dagger$, then up to a normalisation constant, we will finish with the same state each time. This follows our intuition: if we start with ten black marbles, and then add one navy marble to the group and then one mauve marble; we will have the same set of marbles as if we added the mauve one first and then the navy one. Some features concerning the way the marbles are arranged might depend on which one we add first, but what particles we have cannot depend on in what order the particles are added. Thus,

$$\hat{a}_n^\dagger \hat{a}_m^\dagger |1\rangle = \lambda_{nm} \hat{a}_m^\dagger \hat{a}_n^\dagger |1\rangle , \tag{641.1}$$

where λ_{nm} is some complex number. Equally, we find,

$$\hat{a}_m^\dagger \hat{a}_n^\dagger |1\rangle = \lambda_{mn} \hat{a}_n^\dagger \hat{a}_m^\dagger |1\rangle , \tag{641.2}$$

and from this we conclude that $\lambda_{mn}\lambda_{nm} = 1$. Now let us rotate our basis, and in the new basis the creation operators will be $\hat{b}_n^\dagger = R_{nm}\hat{a}_m^\dagger$. We see that,

$$0 = \hat{a}_n^\dagger \hat{a}_m^\dagger |1\rangle - \lambda_{nm} \hat{a}_m^\dagger \hat{a}_n^\dagger |1\rangle = R_{np} R_{mq} (\hat{b}_p^\dagger \hat{b}_q^\dagger - \lambda_{nm} \hat{b}_q^\dagger \hat{b}_p^\dagger) |1\rangle . \tag{641.3}$$

We conclude that $\lambda_{nm} = \lambda_{pq}$. But equally, we could have written this equation as,

$$0 = \hat{a}_n^\dagger \hat{a}_m^\dagger |1\rangle - \lambda_{nm} \hat{a}_m^\dagger \hat{a}_n^\dagger |1\rangle = R_{nq} R_{mp} (\hat{b}_q^\dagger \hat{b}_p^\dagger - \lambda_{nm} \hat{b}_p^\dagger \hat{b}_q^\dagger) |1\rangle . \tag{641.4}$$

So $\lambda_{nm} = \lambda_{pq} = \lambda_{qp}$. Since $\lambda_{pq}\lambda_{qp} = 1$, we conclude that λ_{nm} is either 1 or -1. This means there are two classes of particles, those whose creation operators satisfy ($\lambda_{nm} = -1$)

$$(\hat{a}_m^\dagger \hat{a}_n^\dagger + \hat{a}_n^\dagger \hat{a}_m^\dagger) |1\rangle = 0, \tag{641.5}$$

and those whose creation operators satisfy ($\lambda_{nm} = 1$)

$$(\hat{a}_m^\dagger \hat{a}_n^\dagger - \hat{a}_n^\dagger \hat{a}_m^\dagger) |1\rangle = 0. \tag{641.6}$$

The first of these are known as Fermions, the second Bosons. All particles must fall into one of these two classes. Since we haven't specified the state $|1\rangle$, the relationship between the creation operators is quite general, and doesn't depend on the state of the universe the operators are applied to. Note that if we set $n = m$ in equation (639.5), then we find $\hat{a}_n^\dagger \hat{a}_n^\dagger = 0$, i.e. we can't have two Fermions in the same state at the same time (the Pauli exclusion principle). Also, whenever we swap the order of a pair of Fermion creation operators, we pick up a minus sign in the amplitude. If it seems strange that a non-zero quantity can square to zero, that is because you are thinking in terms of pure numbers rather than operators. It is true that for a pure number, if it squares to zero it must be zero; but different rules apply for operators (such as matrices), and there is no difficulty (for example, it is easy to construct a matrix which is non-zero but squares to zero).

Thus Fermion creation operators anti-commute with each other,

$$\{\hat{a}_m^\dagger, \hat{a}_n^\dagger\} = \hat{a}_m^\dagger \hat{a}_n^\dagger + \hat{a}_n^\dagger \hat{a}_m^\dagger = 0, \tag{642.1}$$

while Boson creation operators commute with each other,

$$[\hat{a}_m^\dagger, \hat{a}_n^\dagger] = \hat{a}_m^\dagger \hat{a}_n^\dagger - \hat{a}_n^\dagger \hat{a}_m^\dagger = 0, \tag{642.2}$$

■

Now we have the notation sorted out, we can turn to physics. I will highlight two questions which physicists are interested in answering[3]: 1) If the universe is in a state $|\psi(t_0)\rangle$ at time t_0 what is the probability that it will be in a state $|\psi(t_1)\rangle$ at time t_1? 2) How can I relate these states with the measurements I actually perform in an experiment? I'll start by addressing the second question.

First of all, we have to ask whether we should consider Fermions or Bosons. This is one of those areas where in principle (I believe) God had a free choice concerning how to construct the universe. To be as general as possible, I shall include both types of field, but I have to study them one at a time. I will start out by considering Fermions, so all the states $|\psi\rangle$ for the moment refer to Fermion fields. Later on, I will add some Bosons into the mix as well.

CONJECTURE 15.3.11. *There is an operator which can be applied to a state vector and a comparison vector which corresponds to the average value of measurements of an observable.*

Proof:
Let us suppose that there is a property of matter σ which can take values $\sigma_1, \sigma_2, \ldots$ (I have written these as discrete values, but the notation can be generalised to a continuous property), with $\left|\psi_i^{(1)}\right\rangle$ representing the states where the value is σ_1, $\left|\psi_i^{(2)}\right\rangle$ the states where the property has value σ_2 and so on, and $\left|\psi_i^{(0)}\right\rangle$ representing states where no particle corresponding to the property is actualised and the property can't be defined. In each case, the index i runs over every possible state which possesses that particular property. This set of states is complete and orthogonal (since a particle can't simultaneously be in two different states), so we can write a general universal state $|\psi\rangle$ as a combination of these basis vectors,

$$|\psi\rangle = \sum_i c_i^{(0)} \left|\psi_i^{(0)}\right\rangle + \sum_j c_j^{(1)} \left|\psi_j^{(1)}\right\rangle + \sum_k c_k^{(2)} \left|\psi_k^{(2)}\right\rangle + \ldots, \tag{642.3}$$

so here we take each possible state, multiply it by the likelihood that the system is in that state, and add them all up. The average value of the observable is thus, using the definition in (15.3.6) and the basic rules of probability (the average value is the probability that it is in a state which gives σ_1 times σ_1 plus the probability that it is in a state which gives σ_2 times σ_2 and so on),

$$\langle\sigma\rangle = \sum_i {c_i^{(0)}}^\dagger 0 c_i^{(0)} + \sum_j {c_j^{(1)}}^\dagger \sigma_j c_j^{(1)} + \sum_j {c_k^{(2)}}^\dagger \sigma_k c_k^{(2)} + \ldots. \tag{642.4}$$

This can be re-written as

$$\langle\sigma\rangle = \langle\psi| \left(\sum_i \left|\psi_i^{(0)}\right\rangle 0 \left\langle\psi_i^{(0)}\right| + \sum_j \left|\psi_j^{(1)}\right\rangle \sigma_1 \left\langle\psi_j^{(1)}\right| + \sum_k \left|\psi_k^{(2)}\right\rangle \sigma_2 \left\langle\psi_k^{(2)}\right| + \ldots \right) |\psi\rangle . \tag{642.5}$$

[3]One can also ask about the properties of various types of materials, but we need the fundamental theory in place before that can be addressed. Properties are related to our measurements; with the fundamental theory in place, we can then consider how to extend it to compute the properties of compound substances.

$\langle\sigma\rangle$ represents the average value after the measurement is taken a large number of times, and thus

$$\hat{\sigma} = \left(\sum_i \left|\psi_i^{(0)}\right\rangle 0 \left\langle\psi_i^{(0)}\right| + \sum_j \left|\psi_j^{(1)}\right\rangle \sigma_1 \left\langle\psi_j^{(1)}\right| + \sum_k \left|\psi_k^{(2)}\right\rangle \sigma_2 \left\langle\psi_k^{(2)}\right| + \ldots \right) \tag{643.1}$$

is an operator which corresponds to the measurement of that observable property. Since we have just constructed an operator that allows us to compute the expectation value, we can say that such an operator exists.

We can also, for example, given a state $|\psi\rangle$ calculate the probability that we will measure the property to have a value σ_1:

$$\begin{aligned} \sum_n c_n^{(1)\dagger} c_n^{(1)} = \langle\psi| & \left(\sum_i \left|\psi_i^{(0)}\right\rangle 0 \left\langle\psi_i^{(0)}\right| + \sum_j \left|\psi_j^{(1)}\right\rangle 1 \left\langle\psi_j^{(1)}\right| + \sum_k \left|\psi_k^{(2)}\right\rangle 0 \left\langle\psi_j^{(2)}\right| + \ldots \right) |\psi\rangle \\ &= P(\sigma = \sigma_1) \end{aligned} \tag{643.2}$$

■

Discussion:

It is worth making an observation here. The uncertainty comes because we can't predict what God is going to do (or what God has done unless we make a measurement). The universe is not in reality a probability amplitude; it is either in one state or another. Suppose that we start with the universe in a state $|\psi\rangle$, and we measure σ and find it to be σ_1. Now the probability that the particle has the observable σ_1 is not as given above, it is 1. The probability given above is the probability of what we will find before we take the measurement; after we have taken the measurement the probability is either going to be 1 or 0. We still don't know what state the universe is in; we only know that it is in one of the states where the particle is in a state where the result of the measurement is σ_1. Therefore, after the measurement, the amplitude for the state of the universe is given by

$$\frac{\sum_n \left|\psi_n^{(1)}\right\rangle \langle\psi_n^{(1)}|\psi\rangle}{\sqrt{\sum_n \langle\psi|\psi_n^{(1)}\rangle\langle\psi_n^{(1)}|\psi\rangle}}. \tag{643.3}$$

To ensure that the probability amplitude remains normalised, we have had to introduce the square root in the denominator. Some may interpret (641.3) as saying that our measurement has changed the state of the universe. However, in this model, $|\psi\rangle$ does not represent the state of the universe but only our knowledge of it; and there is nothing weird in saying that our knowledge of the state of the universe changes after a measurement.

In equation (641.1), the operator $\hat{\sigma}$ that represents the process of measuring the value of the observable σ is given as the sum of various terms such as $\left|\psi^{(i)}\right\rangle \sigma_i \left\langle\psi^{(i)}\right|$. But am I going too far too fast by writing it in this form? Is this the most general expression? In practice, it is. We have already seen that there is a symmetry between the state vectors and comparison vectors for a given state. This symmetry will extend to the operators used to measure observables – we want the same result if we apply the operator to a comparison vector and then compare that against the state vector, as when we apply the operator to the state vector and then use the comparison vector, because in both cases we are essentially doing the same thing. Mathematically, the operators which have this property are known as Hermitian operators; and one of the features of a Hermitian operator is that it can always be expressed in a diagonal form with real eigenvalues – by which I mean the same form as the expansion from (641.1). So we can always choose some basis where the operator is represented by something similar to (641.1). Physically, the eigenvectors represent states which have a definite value of the property σ. Mathematics assures us that this representation of the

operator is always possible. I will call this representation a diagonal representation, using an analogy from the mathematics of matrices (which are a form of operator).

However, although every operator representing some observable property can be expressed in a diagonal form; that diagonal form only exists with one particular choice of basis for the state vectors. Choose a different set of basis vectors, and the operator need not be diagonal. This means that a state in one of those basis vectors will not have a definite value for the operator. How can this be? The vectors parametrise our uncertainty. So if the state vector is given by $\left|\psi^{(1)}\right\rangle$, we are saying that we know that when we take the measurement we will get a result σ_1. If the state vector is given by a superposition of $\left|\psi^{(1)}\right\rangle$ and $\left|\psi^{(2)}\right\rangle$, then we are saying we are not sure what the outcome of the experiment will be; it could be either σ_1 or σ_2. Of course, we could find out easily – by performing the experiment – and once we have done so we can be sure that it is in one or the other of those two states.

So how do we know that every observable can be represented diagonally with the same basis vectors? The simple answer is that there is no such assurance. So for our observable σ, the basis vectors which represent the states with a definite value for the associated property are $\left|\psi_j^{(i)}\right\rangle$. Some other observable Σ, might well be diagonalised by a different set of basis vectors $\left|\Psi_j^{(i)}\right\rangle$. There is no need for these basis vectors to be identical. $\left|\psi_j^{(i)}\right\rangle$ is not a well defined state for the observable Σ; if a state has a definite value for the observable or property σ, it cannot simultaneously have a definite value of the property Σ. This is a bit like saying that if we know that a man has blond hair, then we cannot specify even in principle his eye colour; if we know that a man has blue eyes, then we cannot specify his hair colour. We cannot look at the hair and the eyes together. We look at the hair first – blond. Then we look at the eyes – blue. But the point is that if this man lived in the quantum realm, the state with blue eyes would not be a state of definite hair colour. If a quantum man has blue eyes, then his hair colour is always undetermined. So we go back to look at his hair again, it need not be blond. Is there a contradiction here? No, because when we measure his eye colour we change our knowledge of his state and force him into a state of indefinite hair colour. The contradiction would be if the man in the same state simultaneously had blond and dark hair; but there is nothing logically stopping a man having blond hair in one potentia and dark hair in another. The very act of measurement thus changes the state vector representing the being.

Is this weird? Yes, but it only seems so only because our minds are prejudiced by living at a scale where quantum effects are to all practical purposes too small to be noticed. There are two ways of looking at this. When we look at a man's hair colour (a physical man this time, not the quantum man of the previous paragraph), we do so by bouncing a light photon off him. For something as big as a man one photon of light doesn't really make any difference, so for all practical purposes he remains unchanged. However, for something as small as an electron, if we bounce a photon of light off it, then it will change its state. By the very act of looking at it, we shift things from one potentia to another. When we measure the spin of a particle, the experiment itself forces it into a definite spin state; and therefore out of a definite state for some other observable. We change the state of the particle by the very process of measuring. Secondly, we must remember that these $|\psi\rangle$ vectors are used to parametrise our uncertainty given God's freedom to change reality in any way He sees fit. They represent our knowledge of the universe at some particular moment in time. Our knowledge is certainly changed when we make a measurement; so clearly the $|\psi\rangle$ states will change as well. God is perfectly at liberty to change the state of the universe between measurements – as long as such changes are logically consistent and consistent with God's nature.

There is nothing in my notation which requires that all the operators should be all diagonalisable at the same time. It is just a matter of choosing different sets of basis vectors,

which we are free to do. This weird aspect of quantum physics is a necessary consequence of the mathematical formalism required to describe an indeterminate universe. □

The final ingredient we will need is the single particle projection operator, $|\psi_n\rangle\langle\psi_n|$. So, for example, ψ_n might be the state describing a particular particle in a particular condition (for example with one specified spin state, at one spatial location, and so on - each possible value of all the parameters which describe the particle requires a different state vector $|\psi_n\rangle$. So when we move the particle from one location to another, we change it from a state $|\psi_n\rangle$ to a different state $|\psi_{n'}\rangle$). This projection operator acts on a general state $|\psi\rangle$ describing a system of multiple particles which may or may not include the particle described by $|\psi_n\rangle$, and $|\psi_n\rangle\langle\psi_n|\psi\rangle$ returns 0 if ψ does not contain the particle denoted by ψ_n and $|\psi_n\rangle$ (possibly multiplied by some complex phase) if $|\psi\rangle$ does contain that particle in the given state. Similarly, naively $\langle\psi|\psi_n\rangle\langle\psi_n|\psi\rangle$ gives 1 if the state $|\psi\rangle$ includes the particle $|\psi_n\rangle$ and 0 if it excludes it. This is a little bit naive because $|\psi\rangle$ might contain the particle in an ill defined spin state (with respect to the direction measured by s). The easiest way of eliminating this ambiguity is to sum over all possible spin states. For example, a spin half particle can have two spin states $|\psi^s\rangle$ where s takes the values $\pm\frac{1}{2}$, and thus $\sum_s\langle\psi|\psi_n^s\rangle\langle\psi_n^s|\psi\rangle$ indeed returns 1 if the state ψ contains the particle ψ_n in any state, and 0 if it does not. The sum over all these single particle projection operators is just the identity operator, since this returns the state the universe was in if the universe is in any possible state and 0 if it is in no possible state. But the universe must be in some possible state, so the output of this operator is the same as its input, so this is just the identity operator.

CONJECTURE 15.3.12. *There exists an operator which acts on a state vector which describes how God evolves the universe over time.*

Proof:
Although we cannot say that the universe will evolve from one given state to another given state because of God's free will, nonetheless we can say that given an initial state there is a certain likelihood that God will choose to evolve into a particular final state. What that likelihood is we are not yet in a position to say, but there must be some likelihood for each state, since likelihood is just a mathematical measure of uncertainty; whenever there is uncertainty a likelihood can in principle be assigned to it. Our notation of state vectors is set up to deal with this situation. What this means is that given a state at time t_0, $|\psi(t_0)\rangle$ (which corresponds to a given basis vector), we can say that God will evolve the universe into a state vector at time t_1 $|\psi(t_1)\rangle$ (which is a linear supposition of basis vectors, with the coefficients giving the likelihood of those particular states being actualised). ■

Discussion:
We want to be able to calculate the state at time t_1. To do this, I will split the time between t_1 and t_0 into a large number of discrete steps of duration δt; at the end of the argument I will take the limit that $\delta t \to 0$. Now, at time $t_0 + \delta t$, we know that God is free to destroy the state of the universe, and put anything else in its place. Therefore, His action to evolve the universe from this initial state to some final state $|\psi(t_0 + \delta t)\rangle$ can be represented by the operator

$$T_{0,\delta t} = \sum_m |\psi_m\rangle\, e^{i\frac{1}{\hbar}\hat{\mathcal{L}}(m,\psi(t_0),t_0)\delta t}\,\langle\psi(t_0)|\,. \tag{645.1}$$

We can also introduce an operator which evolves the universe backwards in time, and I will denote this as $T^\dagger_{0,\delta t}$. So the operator $T_{0,\delta t}$ destroys the state of the universe at time t_0, and then creates any other possible state of the universe each with some particular amplitude. The function $\hat{\mathcal{L}}$ (which at the moment we keep unspecified to keep things as general as

possible) represents the likelihood that God will move the universe from state $|\psi(t_0)\rangle$ to a final state $|\psi_m\rangle$. $\hbar$ is just some constant, which I have included to ensure that $\frac{1}{\hbar}\hat{\mathcal{L}}\delta t$ remains dimensionless (i.e. a pure number, rather than depending on any units such as length, time or energy). In principle we could absorb this constant into the function $\hat{\mathcal{L}}$, but in practice it will be more useful to keep it explicitly present. This likelihood will obviously depend on the initial state of the universe, the time, and which of the final states we are moving to; but it cannot depend on anything else, since (at this level of the discussion, where we are neglecting the effect of God's general will and purpose) those are the only circumstances which might affect God's decision about which final state to choose. Thus we can write,

$$|\psi(t_0+\delta t)\rangle \propto \sum_m |\psi_m\rangle\, e^{i\frac{1}{\hbar}\hat{\mathcal{L}}(m,\psi(t_0),t_0)\delta t}. \tag{646.1}$$

The constant of proportionality is determined by the need to keep $\langle\psi(t_0+\delta t)|\psi(t_0+\delta t)\rangle = 1$, but I'll fill this in at the end of the calculation. Now, let's consider the next time slice. Again, God is perfectly free to choose any final state, only with a likelihood that depends on the initial, final states and possibly the time and time difference. The operator to move from one state to the next is given by,

$$\sum_{nm} |\psi_n\rangle\, e^{i\frac{1}{\hbar}\hat{\mathcal{L}}(n,m,t_0+\delta t)\delta t}\, \langle\psi_m|\,. \tag{646.2}$$

This again is the most general expression we can provide. The next state is

$$|\psi(t_0+2\delta t)\rangle \propto \sum_{nm} |\psi_n\rangle\, e^{i\frac{1}{\hbar}\hat{\mathcal{L}}(n,m,t_0+\delta t)\delta t} e^{i\frac{1}{\hbar}\hat{\mathcal{L}}(m,\psi(t_0),t_0)\delta t}. \tag{646.3}$$

Now let us study what this means. Each individual choice of n and m (and so on when we extend this to arbitrary times) represents one particular way in which God can evolve the universe. I will call this one possible path. We then sum over n and m: this means that the final state after applying the time evolution operator represents the sum of all possible ways in which God could evolve the universe, each weighted by this exponential factor.

Why an exponential factor? The reason is straight forward: for each path, the universe ends in one particular state. We start with an initial state $\psi(t_0)$ such that $\langle\psi(t_0)|\psi(t_0)\rangle = 1$. If we select one particular path (that is, one particular choice of n, m, and so on for each individual time step), and we apply the time evolution operator enough times, then we end up with a state $|\psi_{\text{Path}}(t)\rangle = T^{\text{Path}}_{0,t}\,|\psi(t_0)\rangle$, where

$$T^{\text{Path}_{0,t}} = \ldots e^{i\frac{1}{\hbar}\hat{\mathcal{L}}(n,m,t_0+\delta t)\delta t} e^{i\frac{1}{\hbar}\hat{\mathcal{L}}(m,\psi(t_0),t_0)\delta t}, \tag{646.4}$$

without the sum over $m, n, \ldots$, but only one definite route or path through space. Now, the universe must be in a definite state at the end of this path. Therefore $\langle\psi_{\text{Path}}(t)|\psi_{\text{Path}}(t)\rangle = 1$ which in turn implies that $T^{\text{Path}}_{0,t}$ preserves the length of the state $|\psi_{\text{Path}}(t)\rangle$. The class of operators which preserve the length of any complex vector they are applied to are known as unitary operators, and every unitary operator can be written as the complex exponential of the imaginary number i multiplied by a Hermitian operator. This in turn requires that $\hat{\mathcal{L}}$ is a Hermitian operator, i.e. $\hat{\mathcal{L}}(n,m,t) = \hat{\mathcal{L}}(m,n,t)^\dagger$. This is a natural way of writing a unitary operator. The second aspect of this operator is that we have made the exponent proportional to δt. At $\delta t = 0$, we obviously want the update to leave the state unchanged, i.e. the update operator should be the identity operator. We want the universe to be rational (premise 9), which means that the state of the universe at one moment in time should resemble the state of the universe at the next moment in time; there are no sudden discontinuities or jumps in the function used to describe the evolution of the universe. This means that we need the exponent I have written as $\frac{1}{\hbar}\hat{\mathcal{L}}\delta t$ to be proportional to some positive power of δt. Furthermore, if the probability of God leaving nature in the same state for one time step is $1-x$ (for some small x), the the probability that nature is left in the same state for two time steps should be $(1-x)^2 \sim (1-2x) + O(x^2)$. Comparing this to

the probabilities extracted from our expression for the likelihood implies that our exponent should be proportional to the number of time steps, or proportional to δt. We can neglect the δt^2 terms if δt is small enough (and, at the end, we will take the limit $\delta t \to 0$, so it is going to be small enough); they will be too small to affect anything.

So far, we have focused on the idea that God will evolve the system according to one particular path; we have been neglecting the sum over n and m in equation (644.3). However, we don't know which path God will choose. Thus our overall state vector which gives the likelihood of each possible state at some specified future time will involve a sum over all these paths. Adding in the normalisation constant, we find,

$$|\psi(t_1) \setminus \psi(t_0)\rangle = \frac{\sum_{\text{Path}} T^{\text{Path}}_{t_0,t_1} |\psi(t_0)\rangle}{\sqrt{\left(\sum_{\text{Path}'} \langle\psi(t_0)| T^{\text{Path}'\dagger}_{t_0,t_1}\right)\left(\sum_{\text{Path}} T^{\text{Path}}_{t_0,t_1} |\psi(t_0)\rangle\right)}}. \tag{647.1}$$

The notation $|\psi(t_1) \setminus \psi(t_0)\rangle$ means the state vector for time t_1 given that it was in the particular state vector $|\psi(t_0)\rangle$ at time t_0. The square root in the denominator is to normalise this so that $\langle\psi(t_1) \setminus \psi(t_0)|\psi(t_1) \setminus \psi(t_0)\rangle = 1$. Note that the time evolution operator acts on the state vectors themselves, rather than the probabilities. This is a natural consequence of our notation for these physical states.

I shall call $e^{i\frac{1}{\hbar}\hat{\mathcal{L}}}$, which represents the probability amplitude of moving from one state to another, the transfer operator. □

CONJECTURE 15.3.13. *It is possible to count the number of ways in which God can evolve the universe from a given starting state to a given finishing state.*

Proof:
Now let us consider the universe at a time t_0 in a given reference frame, and ask ourselves what state the universe will be in at a later time which I will denote as t_∞. We cannot compare states at different times without using the time evolution operator to evolve the first of these to the later time – just as we cannot compare two things unless we can view them simultaneously.

We have shown that the state vector describing how the system evolves into the future is given by

$$|\psi(t_\infty) \setminus \psi(t_0)\rangle = \frac{\sum_{\text{Path}} T^{\text{Path}}_{t_0,t_\infty} |\psi(t_0)\rangle}{\sqrt{\left(\sum_{\text{Path}'} \langle\psi(t_0)| T^{\text{Path}'\dagger}_{t_0,t_\infty}\right)\left(\sum_{\text{Path}} T^{\text{Path}}_{t_0,t_\infty} |\psi(t_0)\rangle\right)}}. \tag{647.2}$$

$|\psi(t_\infty) \setminus \psi(t_0)\rangle$ represents our knowledge of the state of the universe at time t_∞ conditional upon it being in the specified state at time t_0. But what we are most interested in is what is the probability that it will be in some specific state $|\psi_*(t_\infty)\rangle$ at that time. For example, if the process being described here was the toss of a coin, then $|\psi(t_0)\rangle$ represents how we hold the coin before it is tossed, $|\psi(t_\infty) \setminus \psi(t_0)\rangle$ the likelihood describing the possible states in which the coin could be in given our uncertainty over how hard it will be thrown, and $|\psi_*\rangle$ represents the state where the coin comes up heads. What we want to be asking is what is the probability that it will come up heads given the initial state and our parametrisation of the uncertainty concerning the throwing method.

$$\begin{aligned} P(\psi_*(t_\infty) \setminus t_0) =& \langle\psi(t_\infty) \setminus \psi(t_0)|\psi_*(t_\infty)\rangle\langle\psi_*(t_\infty)|\psi(t_\infty) \setminus \psi(t_0)\rangle \\ =& \frac{\langle\psi(t_0)| \sum_{\text{Path}'} T^{\text{Path}'\dagger}_{t_0,t_\infty} |\psi_*(t_\infty)\rangle \langle\psi_*(t_\infty)| \sum_{\text{Path}} T^{\text{Path}}_{t_0,t_\infty} |\psi(t_0)\rangle}{\langle\psi(t_0)| \sum_{\text{Path}'} T^{\text{Path}'\dagger}_{t_0,t_\infty} \sum_n |\psi_n(t_\infty)\rangle \langle\psi_n(t_\infty)| \sum_{\text{Path}} T^{\text{Path}}_{t_0,t_\infty} |\psi(t_0)\rangle}, \end{aligned} \tag{647.3}$$

where I have used the completeness relation,

$$1 = \sum_n |\psi_n(t_\infty)\rangle \langle\psi_n(t_\infty)| \,, \tag{648.1}$$

and the sum is over every possible state of the universe at time t_∞. The classical definition of probability is that the probability of an event A is the number of ways in which that outcome A is achieved divided by the total number of ways any outcome can be achieved. Equation (645.3) has the same form. Each $|\psi_n\rangle$ state is one particular possible outcome at time t_∞. $|\psi_*(t_\infty)\rangle$ is one of these states; the particular one we happen to be interested in. We can thus interpret the denominator as being proportional to the number of ways in which the system could evolve, and the numerator as proportional to – with the same constant of proportionality – the number of ways in which the system can evolve to reach that particular outcome. That means that the number of ways in which God can reach the final state, $|\psi_*\rangle$, from the initial state, $|\psi(t_0)\rangle$, is proportional to the numerator of equation (645.3),

$$Z = \sum_{\text{Path},\text{Path}'} \langle\psi(t_0)| {T^{\text{Path}'}_{t_0,t_\infty}}^\dagger |\psi_*(t_\infty)\rangle \quad \langle\psi_*(t_\infty)| T^{\text{Path}}_{t_0,t_\infty} |\psi(t_0)\rangle \,. \tag{648.2}$$

What is the constant of proportionality? Well, let us consider the case when there is only one possible path from $|\psi(t_0)\rangle$ to $|\psi_*(t_\infty)\rangle$. In this case, $T^{\text{Path}}_{t_0,t_\infty} |\psi(t_0)\rangle$ will be $|\psi_*(t_\infty)\rangle$ times some complex constant (to preserve the normalisation of $|\psi_*(t_\infty)\rangle$). Since $T^{\text{Path}}_{t_0,t_\infty}$ is a unitary operator, this constant has to be a complex phase, $e^{i\theta}$, so $\langle\psi_*(t_\infty)| T^{\text{Path}}_{t_0,t_\infty} |\psi(t_0)\rangle = e^{i\theta}$. This means that $\langle\psi(t_0)| {T^{\text{Path}}_{t_0,t_\infty}}^\dagger |\psi_*(t_\infty)\rangle = e^{-i\theta}$, since it is just the complex conjugate, and Z is just these two quantities multiplied together, i.e. 1. The number of ways in which God can evolve the system is (in this case) just one, so Z counts the number of ways in which God could evolve the system to get this particular outcome. In this case, the constant of proportionality between Z and the number of ways is just 1. However, this constant is independent of how many ways we have to reach the final state: it has to be the same in each choice of final state. Therefore Z itself is the number of ways in which God can evolve the system to reach the final state.

Intuitively, this seems a little bit odd. We can split Z into two portions, one describing a path forwards in time from from t_0 to t_∞, and one describing a path backwards in time from t_∞ to t_0. We might expect in counting the number of paths to get from A to B we should use the same path going forwards and backwards. This would lead to an expression

$$Z' = \sum_{\text{Path}} \langle\psi(t_0)| T^{\text{Path}}_{t_0,t_\infty} |\psi_*(t_\infty)\rangle \quad \langle\psi_*(t_\infty)| {T^{\text{Path}}_{t_0,t_\infty}}^\dagger |\psi(t_0)\rangle \,. \tag{648.3}$$

This seems to count how many paths can go from the state $|\psi(t_0)\rangle$ to the state $|\psi_*(t_\infty)\rangle$. For every path that finishes in the right state, the projector $|\psi_*(t_\infty)\rangle \langle\psi_*(t_\infty)|$ gives 1 multiplied by some complex phase; if it finishes in the wrong state it gives 0. In the sum over all possible paths, this factor picks out those which finish in the right state, and we count the number of paths.

But a count of the number of ways is a frequency, which can be derived from the probability (since it satisfies the same addition rule). $\langle\psi_*(t_\infty)|\psi(t_\infty)\rangle$ is an amplitude. It cannot be directly compared to a frequency, or a count of the number of ways. We have to first convert it into a probability, which means taking the modulus square. This leads to equation (645.3). The presence of the transfer operator, which assigns a complex phase to each path, then allows for interference effects between the different possible paths. It is these interference effects which makes physics more interesting (and indeed allow classical mechanics to emerge from quantum physics in the large distance limit). ■

CONJECTURE 15.3.14. *We can compute the probability that given a initial state of the universe, there will be a state* $|\psi_*\rangle$ *at a given time,* t.

Proof:
Consider the quantity

$$I(\psi_n(t)) = \langle\psi(t_0)| \sum_{P_1} T^{P_1}_{t_0,t}{}^\dagger |\psi_n(t)\rangle \langle\psi_n(t)| \sum_{P_2} T^{P_2}_{t,t_\infty}{}^\dagger \sum_k |\psi_k(t_\infty)\rangle$$
$$\langle\psi_k(t_\infty)| \sum_{P_3} T^{P_3}_{t,t_\infty} |\psi_n(t)\rangle \langle\psi_n(t)| \sum_{P_4} T^{P_4}_{t_0,t} |\psi(t_0)\rangle . \quad (649.1)$$

We express this in this way, with an insertion of the single particle projector in both the forward and backwards time evolution parts of the operator, to ensure that the final result is both real and positive. Since we don't know what the final state $|\psi(t_\infty)\rangle$ will be, we sum over every possible final state. We can only compare states at the same time, so we need time evolution operators in both the forwards and backwards directions. Once again, we can see what this operator does. For each possible path this returns 1 if that path means that there is an actual particle ψ_n at time t and it finishes in the final state $\psi(t_\infty)$, and zero otherwise. When we sum over paths, this counts the number ways in which God could evolve the universe in both the forward and backwards time direction independently which contain the particle ψ_n at time t and which finish in the state at the end of the universe. Each path is not given the same weight in this count, but the weight is determined by the function $\hat{\mathcal{L}}$ which we discuss in more detail later. We now consider the ratio,

$$\frac{I(\psi_*(t))}{Z} = \frac{A_1}{A_2} \quad (649.2)$$

$$A_1 = \langle\psi(t_0)| T^{P_1}_{t_0,t}{}^\dagger |\psi_*(t)\rangle \langle\psi_*(t)| T^{P_2}_{t,t_\infty}{}^\dagger \sum_k |\psi_k(t_\infty)\rangle$$
$$\langle\psi_k(t_\infty)| T^{P_3}_{t,t_\infty} |\psi_*(t)\rangle \langle\psi_*(t)| T^{P_4}_{t_0,t} |\psi(t_0)\rangle$$
$$A_2 = \langle\psi(t_0)| T^{P_a}_{t_0,t}{}^\dagger \sum_n |\psi_n(t)\rangle$$
$$\langle\psi_n(t)| T^{P_b}_{t,t_\infty}{}^\dagger \sum_k |\psi_k(t_\infty)\rangle \langle\psi_k(t_\infty)| T^{P_c}_{t,t_\infty} \sum_{n'} |\psi_{n'}(t)\rangle \langle\psi_{n'}(t)| T^{P_d}_{t_0,t} |\psi(t_0)\rangle ,$$

where I have omitted the sum over all possible paths to save space. $\sum_n |\psi_n(t)\rangle \langle\psi_n(t)|$ represents the sum over all possible states at time t. This measures the ratio of the number of ways in which God could evolve the universe which lead to a specific state $|\psi_*\rangle$ at time t divided by the total number of ways in which God could evolve the universe. But this is nothing more than the classical probability of there being a state $|\psi_*\rangle$ at time t given our initial and final conditions. So we can write,

$$P(\psi_*(t) \setminus \psi(t_0), \hat{\mathcal{L}}) = \frac{I(\psi_*(t))}{Z}. \quad (649.3)$$

as the probability that there will be a particle in a state $|\psi_*\rangle$ at time t given the initial condition of the universe and our time evolution operator. ■

CONJECTURE 15.3.15. *We can compute the probability that the universe will evolve from a state* $|\psi_n\rangle$ *at time* t_1 *to* $|\psi_m\rangle$ *at a later time* t_2

Proof:
We can extend this construction to asking about multiple particles, and construct the probability

$$P(\psi_n(t_1), \psi_m(t_2) \setminus \psi(t_0), \hat{\mathcal{L}}) = \frac{I(\psi_n(t_1), \psi_m(t_2))}{Z}, \tag{650.1}$$

where $I(\psi_n(t_1), \psi_m(t_2))$ just inserts an additional projection operator for a particle in state m at time t_2,

$$I(\psi_n(t_1), \psi_m(t_2)) = \langle\psi(t_0)| {T^{P_1}_{t_0,t_1}}^\dagger |\psi_n(t_1)\rangle \langle\psi_n(t_1)| {T^{P_2}_{t_1,t_2}}^\dagger |\psi_m(t_2)\rangle \langle\psi_m(t_2)| {T^{P_3}_{t_2,t_\infty}}^\dagger$$
$$\sum_k |\psi_k(t_\infty)\rangle \langle\psi_k(t_\infty)| T^{P_4}_{t_2,t_\infty} |\psi_m(t_2)\rangle \langle\psi_m(t_2)| T^{P_5}_{t_1,t_2} |\psi_n(t_1)\rangle \langle\psi_n(t_1)| T^{P_6}_{t_0,t_1} |\psi(t_0)\rangle \tag{650.2}$$

where I sum over all the paths $P_1, \ldots, P_6$, and to get the probability we just need to divide this by Z. And so on, for questions concerning any number of insertion states. Each state, of course, describes an arbitrary number of particles. By expanding this construction to an arbitrary number of insertions at distinct times, we can compute the probability that the universe is in a particular state at one time given a partial knowledge of an initial state at some earlier time. ■

Discussion:
We should ask why we choose this construction. Is there some other way of getting a probability we could use instead? Firstly, most other ways of trying to extract a probability don't make sense. A probability has to be a real and positive number, so (for example) the quantity

$$A = \frac{\langle\psi(t_0)| T_{t_0,t} |\psi_n(t)\rangle \langle\psi_n(t)| T_{t,t_\infty} |\psi(t_\infty)\rangle}{\langle\psi(t_0)| T_{t_0,t} \sum_n |\psi_n(t)\rangle \langle\psi_n(t)| T_{t,t_\infty} |\psi(t_\infty)\rangle} \tag{650.3}$$

is some measure of the likelihood of a particle ψ_n existing at time t given our initial conditions, it is not a probability since this number need not be positive. This also follows for most other plausible constructions we might try for the probability.

We can consider what the various parts of equation (648.2) mean. $\langle\psi_k(t_\infty)| T^{P_4}_{t_2,t_\infty} |\psi_m(t_2)\rangle$ represents the likelihood that if we start at a state $|\psi_m\rangle$ at time t_2 we finish with a state $|\psi_k\rangle$ at time t_∞, which represents the end of the universe. But then we sum over all the possible final states. Thus

$$\sum_k \langle\psi_m(t_2)| {T^{P_3}_{t_2,t_\infty}}^\dagger |\psi_k(t_\infty)\rangle \langle\psi_k(t_\infty)| T^{P_4}_{t_2,t_\infty} |\psi_m(t_2)\rangle$$

represents the probability that starting in state ψ_m we finish in any of the possible final states of the universe. But this probability is one, because the universe must become something. So if (648.2) is correct, then the probability will be

$$ZP(\psi_m(t_2) \setminus \psi_n(t_1)\psi(t_0), \hat{\mathcal{L}}) = \langle\psi(t_0)| {T^{P_1}_{t_0,t_1}}^\dagger |\psi_n(t_1)\rangle \langle\psi_n(t_1)| {T^{P_2}_{t_1,t_2}}^\dagger |\psi_m(t_2)\rangle$$
$$\langle\psi_m(t_2)| T^{P_5}_{t_1,t_2} |\psi_n(t_1)\rangle \langle\psi_n(t_1)| T^{P_6}_{t_0,t_1} |\psi(t_0)\rangle . \tag{650.4}$$

Z is the same denominator used to normalise the probability we saw earlier. If we consider a second possible state $\psi_{m'}$, which is orthogonal to ψ_m, then

$$ZP(\psi_{m'}(t_2) \setminus \psi_n(t_1)\psi(t_0), \hat{\mathcal{L}}) = \langle\psi(t_0)| {T^{P_1}_{t_0,t_1}}^\dagger |\psi_n(t_1)\rangle \langle\psi_n(t_1)| {T^{P_2}_{t_1,t_2}}^\dagger |\psi_{m'}(t_2)\rangle$$
$$\langle\psi_{m'}(t_2)| T^{P_5}_{t_1,t_2} |\psi_n(t_1)\rangle \langle\psi_n(t_1)| T^{P_6}_{t_0,t_1} |\psi(t_0)\rangle . \tag{650.5}$$

This means that

$$Z(P(\psi_m(t_2) \setminus \psi_n(t_1)\psi(t_0), \hat{\mathcal{L}}) + P(\psi_{m'}(t_2) \setminus \psi_n(t_1)\psi(t_0), \hat{\mathcal{L}})) =$$
$$\langle\psi(t_0)| {T^{P_1}_{t_0,t_1}}^\dagger |\psi_n(t_1)\rangle \langle\psi_n(t_1)| {T^{P_2}_{t_1,t_2}}^\dagger (|\psi_m(t_2)\rangle \langle\psi_m(t_2)| + |\psi_{m'}(t_2)\rangle \langle\psi_{m'}(t_2)|)$$

$$T^{P_5}_{t_1,t_2} |\psi_n(t_1)\rangle \langle\psi_n(t_1)| T^{P_6}_{t_0,t_1} |\psi(t_0)\rangle . \quad (651.1)$$

But $(|\psi_m(t_2)\rangle \langle\psi_m(t_2)| + |\psi_{m'}(t_2)\rangle \langle\psi_{m'}(t_2)|)$ is just the projection operator for being in either the state $|\psi_m\rangle$ or the state $|\psi_{m'}\rangle$. Thus we have

$$Z(P(\psi_m(t_2) \setminus \psi_n(t_1)\psi(t_0), \hat{\mathcal{L}}) + P(\psi_{m'}(t_2) \setminus \psi_n(t_1)\psi(t_0), \hat{\mathcal{L}})) = ZP(\psi_{m'}(t_2) \cup \psi_m(t_2) \setminus \psi_n(t_1)\psi(t_0), \hat{\mathcal{L}}), \quad (651.2)$$

which is the defining equation of a probability. Thus $\frac{I}{Z}$ indeed represents the probability that given initial states $|\psi(t_0)\rangle$ and $|\psi_n(t_1)\rangle$ and the (as yet undetermined) physical time evolution rules contained within $\hat{\mathcal{L}}$ we should find ourselves in the state $|\psi_m\rangle$ at time t_2. □

15.4. Fermion States

So far, I have just discussed discrete states. This makes the notation somewhat simpler. In reality, however, the states are not discrete but continuous – just as space, time, energy and momentum are continuous variables. It is a small step to get from the discrete to the continuous – we need to change matrices into operators and vectors into functions. But this is something I now have to do. I will also take this chance to switch from discussing state vectors and instead re-introduce the creation and annihilation operators.

A state vector containing various particles with different indices $\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \ldots$ can be written as

$$|\psi(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \ldots)\rangle = \hat{a}^\dagger_{\vec{\mathbf{x}}_1} \hat{a}^\dagger_{\vec{\mathbf{x}}_2} \ldots |0\rangle . \quad (651.3)$$

We have assumed that the universe is described by space and time, and each particle is located in one particular location, and to make it easier when I want to consider Lorentz invariance a little later I have used 4-vectors here rather than 3 vectors. So $\vec{\mathbf{x}}_1$ physically represents the location of the particle and the time it is created. It is a continuous index used to distinguish between different particles. This many particle state vector satisfies the orthonormality conditions, and we can use this to derive the anti-commutation relations between $\hat{a}$ and $\hat{a}^\dagger$.

CONJECTURE 15.4.1. *For a Fermion field,* $\{\hat{a}_{\vec{\mathbf{x}}_1}, \hat{a}^\dagger_{\vec{\mathbf{x}}_2}\} = \delta^3(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2))\hat{\beta}$*, where* $\hat{\beta}$ *is the operator which gives* 1 *for a particle and* -1 *for an anti-particle and* δ *is the Dirac* δ*-function.*

Proof:
Fermions are particles whose creation and annihilation operators anti-commute with each other, so we consider the anti-commutator here rather than the commutator.

The state which describes a particle at location $\vec{\mathbf{x}}_1$ is $|\psi(\vec{\mathbf{x}}_1)\rangle = \hat{a}^\dagger_{\vec{\mathbf{x}}_1} |0\rangle$. The orthonormality condition for continuous states reads $\langle\psi(\vec{\mathbf{x}}_2)|\psi(\vec{\mathbf{x}}_1)\rangle = \delta^3(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2)\hat{\beta}$. I have added the operator $\hat{\beta}$ into this equation because I have looked ahead and know where I want to go with this argument. Eventually, $\hat{\beta}$ will be needed to distinguish particles from anti-particles (an anti-particle is basically a hole or lack in a particle field, and for this reason we want the orthonormality condition to give -1 for an anti-particle rather than 1), but for now we just treat it as some operator put into the equation because we can and it costs us nothing to do so. We can always set $\hat{\beta}$ to be the identity operator later if we find that it isn't needed in practice. Thus, for two states describing single particles created at the same time,

$$\begin{aligned}\langle\psi(\vec{\mathbf{x}}_2)|\psi(\vec{\mathbf{x}}_1)\rangle = \delta^3(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2)\hat{\beta} &= \langle 0| \hat{a}_{\vec{\mathbf{x}}_2} \hat{a}^\dagger_{\vec{\mathbf{x}}_1} |0\rangle \\ &= \langle 0| (\{\hat{a}_{\vec{\mathbf{x}}_2}, \hat{a}^\dagger_{\vec{\mathbf{x}}_1}\} - \hat{a}^\dagger_{\vec{\mathbf{x}}_1} \hat{a}_{\vec{\mathbf{x}}_2}) |0\rangle \\ &= \langle 0| \{\hat{a}_{\vec{\mathbf{x}}_2}, \hat{a}^\dagger_{\vec{\mathbf{x}}_1}\} |0\rangle . \end{aligned} \quad (651.4)$$

I have used $\hat{a}_{\vec{\mathbf{x}}_2}|0\rangle = 0$, i.e. you can't annihilate a particle that isn't there. The Dirac δ-function is used to indicate that if the states describe particles created at the same location – i.e. the states are identical – then this quantity multiplied by an infinitesimal element of volume should give 1 to say that the two states are identical. However, if the two particles are created at different locations – the states are different – then the comparison should give zero. There is no need to use the vacuum state $|0\rangle$ in this equation. For example, if we consider any state $|A\rangle$, containing any number and distribution of particles except it lacks the particle at position $\vec{\mathbf{x}}_2$ (i.e. $\hat{a}_{\vec{\mathbf{x}}_2}|A\rangle = 0$), and consider the state which is A plus a particle at $\vec{\mathbf{x}}_1$, $|\psi(\vec{\mathbf{x}}_1, A)\rangle = \hat{a}^\dagger|A\rangle$, and apply the same reasoning, we find,

$$\langle A|\{\hat{a}_{\vec{\mathbf{x}}_2}, \hat{a}^\dagger_{\vec{\mathbf{x}}_1}\}|A\rangle = \delta^3(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2)\hat{\beta}. \tag{652.1}$$

Since the state A is general and could be anything, we conclude that

$$\{\hat{a}_{\vec{\mathbf{x}}_2}, \hat{a}^\dagger_{\vec{\mathbf{x}}_1}\} = \delta^3(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2)\hat{\beta}. \tag{652.2}$$

■

Discussion:
We can, of course, extend this to examine specified many particle states (where many = n), i.e. $|\psi(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \vec{\mathbf{x}}_3, \ldots)\rangle = \hat{a}^\dagger_{\vec{\mathbf{x}}_1}\hat{a}^\dagger_{\vec{\mathbf{x}}_2}\hat{a}^\dagger_{\vec{\mathbf{x}}_3}\ldots|0\rangle$. The orthonormality condition here is

$$\langle\psi(x_a, x_b, x_c, \ldots)|\psi(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \vec{\mathbf{x}}_3)\rangle = \delta(x_a - \vec{\mathbf{x}}_1)\delta(x_b - \vec{\mathbf{x}}_2)\delta(x_c - \vec{\mathbf{x}}_3)\hat{\beta}^n\ldots - \delta(x_a - \vec{\mathbf{x}}_2)\delta(x_b - \vec{\mathbf{x}}_2)\delta(x_c - \vec{\mathbf{x}}_3)\hat{\beta}^n\ldots + \ldots, \tag{652.3}$$

where we alternate signs as we swap the order of the creation operators.

These minus signs are a bit annoying, and we need some convention to ensure we are consistent. In the state $|\psi(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \vec{\mathbf{x}}_3, \ldots)\rangle$, I will order the creation operators so that they are in the order of increasing time, which fits in well with the path integral formulation.

Earlier, I wrote the creation operator as $\hat{a}^\dagger_{\vec{\mathbf{x}}_1} = \sum_A |\psi(\vec{\mathbf{x}}_1 \cup A)\rangle\langle\psi(A)|$, where the sum over A denotes the sum over all possible states of the universe excluding a particle at location $\vec{\mathbf{x}}_1$, and $\vec{\mathbf{x}}_1 \cup A$ indicates the state you get when you take the state A and add into it another particle at $\vec{\mathbf{x}}_1$. For Fermions, we need to be a little bit more careful, because the creation operators should anti-commute, but this representation doesn't. So it is worth looking at this in a little more detail. To keep things simple, I will restrict myself to a discrete system with two possible Fermions, which I will label as Fermion 1 and Fermion 2. This can be extended easily to any number of particles, and where they are described by a continuous rather than discrete index, but the two particle system is both simple and enough to understand the general idea. I will describe $|\psi_{00}\rangle$ as the state with no particles, $|\psi_{10}\rangle$ as the state with particle 1 in place but not particle 2, $|\psi_{01}\rangle$ as the state with Fermion 2 but not Fermion 1, and $|\psi_{11}\rangle$ as the state where both particles are present. Since there can only be one particle in each state, these four states are complete.

However, before we draw up the expressions for the creation operators, we need to define the $|\psi_{11}\rangle$ state a little more carefully. The order in which we create the particles is important. I will say that $|\psi_{11}\rangle$ is defined assuming that we create Fermion 1 first and Fermion 2 second. If we created them the other way round, then there would be a relative minus sign in front of the state. This particular choice of convention is not so important; what is important is that we consistently use whatever convention we choose throughout the calculation.

So the following are the expressions for the creation and annihilation operators:

$$\hat{a}^\dagger_{F1} = |\psi_{10}\rangle\langle\psi_{00}| - |\psi_{11}\rangle\langle\psi_{01}| \qquad \hat{a}_{F1} = |\psi_{00}\rangle\langle\psi_{10}| - |\psi_{01}\rangle\langle\psi_{11}|$$
$$\hat{a}^\dagger_{F2} = |\psi_{01}\rangle\langle\psi_{00}| + |\psi_{11}\rangle\langle\psi_{10}| \qquad \hat{a}_{F2} = |\psi_{00}\rangle\langle\psi_{01}| + |\psi_{10}\rangle\langle\psi_{11}|. \tag{652.4}$$

That minus sign in the creation operator $\hat{a}^\dagger_{F1}$ is because the particles are created in the wrong order.

So, with this notation, it is fairly easy to check that

$$\hat{a}^\dagger_{F1}\hat{a}_{F2} + \hat{a}_{F2}\hat{a}^\dagger_{F1} = |\psi_{10}\rangle\langle\psi_{00}|\psi_{00}\rangle\langle\psi_{01}| - |\psi_{10}\rangle\langle\psi_{11}|\psi_{11}\rangle\langle\psi_{01}| = 0, \quad (653.1)$$

as we expect. Incidentally, we can also imagine what happens when the indices represent some position and we gradually move the two particles together. As soon as their separation reaches zero and they exist in the same location, the $|\psi_{11}\rangle$ state is forbidden, since we can't have two Fermions of the same type in the same location. This means that the term containing $|\psi_{11}\rangle$ in equation (651.1) is excluded, and we get something non-zero (note that it is positive – this will be important in a moment), in agreement with equation (650.2).

In other words, the expression of creation and annihilation operators for Fermions in terms of the more fundamental states is still valid, but we just have to be careful because various minus signs enter the equation.

This leads us to another thought. I have been discussing the $|\psi_{10}\rangle$ state as though it meant that the particle was present, and $|\psi_{00}\rangle$ as though the particle were absent. This is a nice physical interpretation, but there is a mathematically oddity which we ought to note. There is a great deal of symmetry between the 'absent' and 'present' states. This is, effectively, a two state system, with nothing mathematically distinguishing between the two states. So who are we to say that it is the $|\psi_{00}\rangle$ which is fundamental, and not the $|\psi_{11}\rangle$ state? In other words, why is the basis of our calculation the state with no particles present? When we inserted the minus sign, we assumed that we created particles from the vacuum state $|\psi_{00}\rangle$; but we can achieve the same effect by annihilating particles from the everything present state, $|\psi_{11}\rangle$. Could it not also be the state with all particles present, and the fluctuations we know as particles are subtractions from this foundational state?

So, we can consider what happens when we swap the 1 and 0 indices in the creation and annihilation operators. We get,

$$\begin{aligned} \hat{a}_{A1} &= |\psi_{01}\rangle\langle\psi_{11}| - |\psi_{00}\rangle\langle\psi_{10}| & \hat{a}^\dagger_{A1} &= |\psi_{11}\rangle\langle\psi_{01}| - |\psi_{10}\rangle\langle\psi_{00}| \\ \hat{a}_{A2} &= |\psi_{10}\rangle\langle\psi_{11}| + |\psi_{00}\rangle\langle\psi_{01}| & \hat{a}^\dagger_{A2} &= |\psi_{11}\rangle\langle\psi_{10}| + |\psi_{01}\rangle\langle\psi_{00}|. \end{aligned} \quad (653.2)$$

So the annihilation operator here creates a 'hole' in the fully occupied state of Fermions. This all leads to

$$\hat{a}_{A1}\hat{a}^\dagger_{A2} + \hat{a}^\dagger_{A2}\hat{a}_{A1} = |\psi_{01}\rangle\langle\psi_{11}|\psi_{11}\rangle\langle\psi_{10}| - |\psi_{01}\rangle\langle\psi_{00}|\psi_{00}\rangle\langle\psi_{10}| = 0. \quad (653.3)$$

And one can also check that the other commutation relations are satisfied. Physically, this system is identical to the one we had earlier. We cannot distinguish between particles created from a no particle state, or holes created from a universe where all the Fermion states are fully occupied. The only difference between this and the previous example is the odd minus sign here or there. For example, when we bring the two particles closer together, it is again the $|\psi_{11}\rangle$ term which is forbidden by the exclusion principle, but this time it is the first term in the equation that consequently vanishes. We thus find that for these holes,

$$\{\hat{a}^\dagger_{A\vec{\mathbf{x}}_1}, \hat{a}_{A\vec{\mathbf{x}}_2}\} = -\delta^3(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2) \quad (653.4)$$

while the Fermions themselves satisfied

$$\{\hat{a}^\dagger_{F\vec{\mathbf{x}}_1}, \hat{a}_{F\vec{\mathbf{x}}_2}\} = \delta^3(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2). \quad (653.5)$$

This possible minus sign is accounted for the $\hat{\beta}$ operator which I inserted in my expression for the commutation relation. It gives $+1$ if we are creating particles (the F class of creation operators), and -1 if we are creating holes, or anti-particles (the A class of creation operators). We need to include both possibilities into our construction, and this is easiest by introducing the operator $\hat{\beta}$ so we can use the same equation in both cases.

Why would we expect to see both particles and anti-particles in nature? As far as God is concerned, there really isn't any difference between the two options. A universe consisting entirely of anti-particles would experience God in the same way as a universe consisting entirely of particles (baring differences in electric charge, and the odd minus sign). The operator $\hat{\mathcal{L}}$ which we introduced earlier to describe how God interacts with the universe must therefore describe both the evolution of particles and the evolution of anti-particles, and because there is a symmetry between the two types of universe – they both experience the same God – there must correspondingly be a symmetry between particles and anti-particles in the operator $\hat{\mathcal{L}}$. The easiest way of doing this is to write the creation and annihilation operator as a vector with an even number of components, so in a two component system $\hat{a}_1^\dagger = (\hat{a}_{F1}^\dagger, \hat{a}_{A1}^\dagger)$, or perhaps a four component system (which is in practice what we need) described by $\hat{a}_1^\dagger = (\hat{a}_{F1u}^\dagger, \hat{a}_{F1d}^\dagger, \hat{a}_{A1u}^\dagger, \hat{a}_{A1d}^\dagger)$. In this notation, $\hat{\beta}$ is a diagonal matrix with 1 in its top components, and -1 in its lower components. It is these 4-component creation operators, called spinors, which I will need below. □

15.5. Bell's Theorem

Objection:
What is being advocated here is just a hidden variable theory. Such theories were ruled out by the experimental violations of Bell's inequalities found by Clauser and Aspect. Therefore this interpretation is incorrect. •

Response:
My construction is based around the idea that the quantum mechanical (or, better, field theory) state ψ is a likelihood parametrising our uncertainty concerning the state of the universe. The universe itself is always in a definite state, but we cannot know for certain what that state is unless we have just measured it. Secondly, God's freedom of action implies that the universe is fundamentally indeterminate, so that even if we know what the state of the universe is at one moment in time, we cannot predict with certainty what state it will be in in future times. We can, however, parametrise our uncertainty in terms of likelihoods. This conception obviously has similarities with the hidden variables theories proposed by Einstein, Rosen and Podolski, and disproved by the experimental confirmation of Bell's response to it. Thus I need to discuss whether the differences between my construction and that of Einstein are enough to stymie Bell's and other similar arguments against hidden variable theories.

Bell's inequalities demonstrate that there is no underlying *classical* physical theory behind quantum mechanics. It was formed in response to an objection by Einstein, and is often used to support either the Copenhagen or Bohm's interpretation of quantum mechanics (or field theory; the Copenhagen interpretation applies in both cases) which states that matter at the small scale where quantum mechanics is valid is not in a real state until a measurement is made; for coupled observables (for example, two particles emerging from the decay of another particle forced by conservation laws to be in opposite or the same spin state) this implies that measurements on one particle instantaneously transform its partner into a real state, even though that partner might be a long distance away, contrary to the spirit, if not the letter, of special relativity. This action at a distance in particular made Einstein nervous.

Let us suppose that a Higgs Boson at rest in the laboratory frame, spin zero, decays into two τ-leptons, which are spin $\frac{1}{2}$ particles. These would travel at large speeds in opposite directions towards two detectors, one at position A and one at position B, and the particles must have opposite spin from the conservation of angular momentum. The detectors

measure the spin of the τs along a particular axis. Because the spin operators in different directions don't commute, they can only measure the spin of the particle in one direction at a time.

Suppose that there is an underlying classical dynamics behind quantum physics; the indeterminacy in measurement relates only to the experimental equipment. That is to say that the particles leave the decay process with a definite spin state; only each particle must have opposite spin. It is only when attempting to make a measurement that probabilities and uncertainties enter into the equation; our experimental apparatus forces the particles to be aligned in a particular direction, and this process of alignment depends on some unmeasured variables describing the τ state, which since we don't know what they are appear to us to give a 'random' measurement. In other words, the apparent indeterminacy of quantum mechanics, and that the measurements give discrete values, is simply because we cannot measure all the variables that describe the spin state of the τ and because of the nature of the process of taking a measurement. This should be set against the 'orthodox' quantum mechanical interpretation where the spin of the τ is wholly undetermined until it is measured.

Suppose first of all that the detectors are aligned so they measure spin along the same axis. In this case the measurements at detector A will be the opposite of those at detector B – one of them will record spin $\frac{1}{2}$, and the other spin $-\frac{1}{2}$. Of course, we cannot a priori say which of them is which because we don't know the underlying hidden spin variables of the particle – the probability of detector A recording the spin $\frac{1}{2}$ particle is one in two, and if the experiment is repeated numerous times roughly half of the particles reaching detector A will be recorded as having positive spin, but the results from detector B will be completely correlated with it. And that is the result we expect from both the hidden variables theory and orthodox quantum mechanics.

Now suppose that we rotate detector A so it measures spin aligned along a direction at a small angle θ to the original axis. Now we can expect the results at A and B will become decorrelated – most of the time we will still get the opposite answer at A and B, but sometimes they will record the same spin, because of the slight misalignment between the detectors. In quantum mechanics (and the hidden variables theory can be constructed to give the same result) the probability of detectors A and B recording the same spins is increased to $P(A_{\text{Moved}}) = \sin^2\theta$. Since detector B records the (opposite) result to what we would have received at detector A if we had not rotated it (and, in the hidden variable theory, this is independent of what we do at detector A), this represents the probability that the result of the experiment at detector A has changed as a result of our rotation of the detector. Now suppose that we restore detector A to its original orientation, and rotate detector B by the same angle but in the opposite direction. Once again, the probability that the detectors record the same spin is increased to $P(B_{\text{Moved}}) = \sin^2\theta$. And again, we achieve the same result result regardless of whether we use the hidden variable or quantum mechanical picture.

But now suppose that we rotate both detectors A and B. According to the hidden variables theory, what happens at each detector ought to be independent of what we do with the other one. So the frequency of changed measurements at detector A should remain the same; so should the frequency of recording the same result at detector B. Thus the probability of recording the same result at each detector is

$$P(A_{\text{Moved}} \cup B_{\text{Moved}}) \leq P(A_{\text{Moved}}) + P(B_{\text{Moved}}) = 2\sin^2\theta, \tag{655.1}$$

where the inequality follows from a standard result from probability theory (it is an equality if $A_{\text{Moved}} \cap B_{\text{Moved}} = \emptyset$). However, quantum physics predicts that

$$P(A_{\text{Moved}} \cup B_{\text{Moved}}) = \sin^2 2\theta, \tag{655.2}$$

which for $-45° < \theta < 45°$ (except $\theta = 0$) is greater than the maximum bound suggested by the classical result. Experiment, of course, agrees with the quantum mechanical prediction in violation of the inequality (653.1).

So what has happened here? The derivation of (653.1) seems to just rely on three assumptions: the values of $P(A_{\text{Moved}})$ and $P(B_{\text{Moved}})$ (which are identical in the indeterminate and hidden variable theories); separability, i.e. that the probability $P(A_{\text{Moved}})$ is independent of what we do at detector B and vice-versa; and reality, namely that the results that will later be measured at detectors A and B are determined by the definite state of the particles as they emerge from the decay. If X is the background information, and a and b the hidden variable states of the two particles, then the independence condition can be written as $P(A_{\text{Moved}} \setminus XaB) = P(A_{\text{Moved}} \setminus Xa)$, and a similar expression for B. This is to say that the probability for the result at A only depends on the precise state of the particle hitting the detector, and the initial conditions at the experiment, and not on what is happening at B. The reason why we might believe this to be the case is locality; what happens at one place cannot influence what happens at another place outside its light cone. Since the inequality is violated, one of these premises must be wrong: the first premise can be checked experimentally, and is solid; either the configuration of detector A affects the results at detector B, or something even more weird is happening. One could, of course, construct similar results using other sets of non-commuting measurements.

So what do we make of this? Firstly, we must always remember that the mathematical description is only a representation of reality, and not reality itself. That something weird is happening in the mathematics does not necessarily mean that something weird is happening in reality – it could just be a quirk of the method of abstraction. But there is a more direct answer.

The picture that I have been working under assumes that the particles emerge from the decay in a particular spin state (expressed in terms of a spin operator with a particular orientation and the eigenvalue of that operator). This then proceeds until we apply the detector. A spin state aligned along one direction can, of course, be expressed as a linear superposition of the spin states in other directions; so if the particle emerges with spin along the z direction, then its spin x result and spin y result would be undetermined. This is, of course, different from a classical hidden variables theory, where all the spin states are supposed to be expressed in some way in terms of hidden variables. If we measure spin along a different axis, then God is free to select whichever one of the two results according to the amplitudes implied by the appropriate linear superposition of states when we rotate from one eigenvector basis to another. The reason we can't give definite results in quantum physics is because of God's freedom to act. The amplitudes are (in part) a parameterization of our predictions of how likely it is that God will act in a particular way. But even so, suggesting that the particle emerges from the decay with a real spin state (as opposed to a potential spin state in each direction) surely leads to Bell's doomed inequality, does it not?

Now we should recall conjecture 15.3.8. In quantum physics, probabilities can only be defined if we express all relevant amplitudes in terms of the same basis state, defined by the same operators. The statement $P(A \cup B)$ makes no mathematical sense unless A and B are expressed in terms of the same basis. But in Bell's theorem, $P(A_{\text{Moved}})$ was measured using operators which used one particular representation of the spin operators, and $P(B_{\text{Moved}})$ was measured using a different set of basis operators. Therefore $P(A_{\text{Moved}} \cup B_{\text{Moved}})$ is not a probability (it denies premise 0 of probability, as defined in section 2.2), and the inequality (653.1) consequently does not follow.

So what should the result be? Suppose that the particles are emitted in a definite state $|S_A, S_B\rangle$, whose spin is orientated along an axis $\hat{\vec{\mathbf{e}}}$ (with spin $\frac{1}{2}$ for A and $-\frac{1}{2}$ for B). Detector A is arranged at an angle θ_A to $\hat{\vec{\mathbf{e}}}$ and detector B is arranged at an angle θ_B to $\hat{\vec{\mathbf{e}}}$. We need to now construct a complete, irreducible and orthogonal basis which we can use to express our probabilities. This basis is the set of states $|A_+, B_+\rangle$, $|A_+, B_-\rangle$,

$|A_-, B_+\rangle$, and $|A_-, B_-\rangle$, where A_+ indicates that detector A records the spin in a positive state. We can write $|S_A\rangle$ as a linear superposition of the $A_\pm$ states: $|S_A\rangle = \cos\theta_A\,|A_+\rangle + \sin\theta_A\,|A_-\rangle$, while $|S_B\rangle = -\sin\theta_B\,|B_+\rangle + \cos\theta_B\,|B_-\rangle$ (this is just the appropriate rotation of basis for spin $\frac{1}{2}$ particles). The amplitude for the detectors to both record positive spin is $\langle A_+, B_+|S_A, S_B\rangle = -\cos\theta_A \sin\theta_B$, while the amplitude for the detectors to both record negative spin is $\langle A_-, B_-|S_A, S_B\rangle = \cos\theta_B \sin\theta_A$; the total amplitude for the detectors to record the same spin state is $\sin(\theta_A - \theta_B)$, and the probability is $\sin^2(\theta_A - \theta_B)$. This is the correct result. The basis needed to express the result in terms of a probability depends on the rotation of the two detectors. However, this rotation is fixed by the experimenters; it does not depend on information passing from one particle to another instantaneously when the measurement is made. Of course, complete knowledge of the direction and state of the particle at B allows one to compute the probability that the two detectors achieve the same result, a calculation only dependent on knowledge of the orientation of the detector at A but not the result of the experiment.

Bell's inequalities thus depend on an additional assumption: that the uncertainties in the state of the particles that result from the hidden variables ought to be parametrised using classical probability. They exclude the possibility of a underlying mechanistic physics behind quantum physics where all properties are determined by the state of the particle (and Kant's suggestion that the state of the particle can be given by listing all the properties of the particle is similarly excluded). The particle can be in a definite (e.g. spin) state, but some of its properties undetermined; indeed if it is in a definite spin state then some of its properties must be undetermined. We can and must specify the state of a system without specifying all its properties; the best we can do is express those properties as likelihoods according to the linear superposition principle. Thus a system built on classical mechanism with uncertainties mapped directly to probabilities is ruled out. Instead, even though it is less intuitive (though less intuitive is not the same as being wrong), we have to treat fundamental objects as quantum mechanical states, with uncertainties represented by amplitudes rather than probabilities, and with observables represented by different bases than the one that most naturally describes the particle state left as undetermined linear superposition. However, the two particles can after the decay exist in such a definite quantum state, and thus since it is in such a state we need not deny its reality, and since there need be no spooky communication between the paired particles when the measurement is made there is still separability.

Bell's inequality does rule out what we might call *property realism*, where behind the formulation of quantum physics is a hidden variables theory where all the properties (or measurement results, or future events) of each particle are completely determined. It does not, however, rule out *substance realism*, where there is an underlying structure which can leave certain properties indeterminate. The quantum state is in a particular state in a given basis, but not all properties are defined in that basis. This indeterminacy concerning properties is all Bell's theorem forces us to adopt. Thus Bell's inequality rules out a classical or mechanistic hidden variables theory, such as those used as the basis of much of enlightenment philosophy, particularly Kant's, but does not rule out the picture I am building here.

The other assumption behind Bell's inequality is locality, the idea that interactions always occur at a single point. The usual statement is that Bell's theorem rules out either property realism or locality. Most physicists will not give up locality, and there are good reasons for this, based partly on special relativity, and partly on the fact that the theorem proving the conservation of energy and momentum in quantum physics (but not classical physics) is derived from the locality of the Hamiltonian. If locality is broken, then so is energy conservation; and without energy conservation the key assumption of science, that the world is rational and understandable, that every effect follows a cause, would be violated.

However, we need to be a bit careful when discussing locality. There is the distinction between event and substance causality. The Hamiltonian and the Lorentz invariance of special relativity each constrain substance causality. They control the decay of one particle into another, and messages carried by particles from one part of the universe to another. But Bell's theorem discusses measurement outcomes, or events. The event in question is whether the particle moves into a spin up state or spin down state when it interacts with the detector. Whichever way it goes, there is no violation of substance causality, or the premise that the process of creation and annihilation of particles is local. Now, in my philosophy, quantum events are ultimately driven by God. Before the particle enters the detector, it has the potential to become spin up or spin down, but it is God who actualises one or the other of those potentials. The amplitude for getting a result A depends on the background information, the state of the particle at that detector, and the choices of God. Now God is aware of what is happening at both detectors. Being timeless, it matters little to God which measurement is performed first. So if we denote $G(A, B)$ as representing the role of God in determining the measurement result, and ψ as the amplitude, we are interested in calculating $\psi(A|XaG(A,B))$ and $\psi(B|XbG(A,B))$. These two amplitudes are not independent of each other, so the separability assumption in the derivation of Bell's inequality without requiring any physical breakdown of locality. There is no natural non-local information flow, but the result of the measurement is determined in part by something supernatural. Now God is about as non-local as it is possible to get, and if events are driven by God, then we should not be surprised if there is an otherwise inexplicable correlation between two different spatially separated events. Thus demanding that substance causality is local satisfies the requirements of relativity and quantum physics, while allowing event causality (the indeterminate part of quantum physics) to violate locality satisfies the demands of Bell's inequalities. The distinction between substance and event causality allows us to understand these otherwise contradictory requirements. ○

15.6. Quantum Electrodynamics

As discussed earlier, an annihilation operator acting on the vacuum state gives zero, $\hat{a}_{\vec{\mathbf{x}}}|0\rangle = 0$, since you can't annihilate something that isn't there. Similarly, $\langle 0|\,\hat{a}^\dagger_{\vec{\mathbf{x}}} = 0$. When we have two operators sandwiched between some vacuum states, then the tactics we can use to simplify the expression is to anti-commute the operators through each other so that all the creation operators are on the left, and the annihilation operators on the right of the expression.

$$\begin{aligned}\langle 0|\,\hat{a}_{\vec{\mathbf{x}}}\hat{a}^\dagger_{\vec{\mathbf{y}}}|0\rangle &= \langle 0|\left(-\hat{a}^\dagger_{\vec{\mathbf{y}}}\hat{a}_{\vec{\mathbf{x}}} + \{\hat{a}_{\vec{\mathbf{x}}}, \hat{a}^\dagger_{\vec{\mathbf{y}}}\}\right)|0\rangle \\ &= \delta^3(\vec{\mathbf{x}} - \vec{\mathbf{y}})\hat{\beta} \end{aligned} \tag{658.1}$$

The same basic idea applies when we compare states with two or more particles: we anti-commute the creation operators to the left of the state, leaving us with a set of delta functions, which effectively mean that only the bracket between two states describing the same particles is non-zero. What if the comparison vector describes a different number of particles to the state vector (counting particles as 1 particle and anti-particles as -1 particles)? Here we also get zero. Suppose, for example, the state (the creation operators on the right of the expression) describes one more particle than the comparison (the annihilation operators on the left). In this case, we anti-commute creation operators through annihilation operators, destroying one of each in the process. But since we have more creation operators than annihilation operators, there will be at least one operator left over when we complete the process. And a creation operator sandwiched between two vacuum states, $\langle 0|\,\hat{a}^\dagger_{\vec{\mathbf{x}}}|0\rangle$, gives zero.

The transfer operator, $e^{\frac{i}{\hbar}\hat{\mathcal{L}}\delta t}$, physically represents the destruction of one set of particles and the creation of another. At each moment of time, God is free to destroy the universe and create a new one however He pleases, or to leave the universe in the same state. Movement of a particle from one state to another is equivalent to destroying the original state and creating the final one. Thus the operator $\hat{\mathcal{L}}$ must consist of some number of creation operators multiplied by some number of annihilation operators, perhaps multiplied by some function of either the location or the relative locations of the particles. Each term within it can therefore be written as something like

$$\hat{\mathcal{L}} = \ldots + \hat{a}^\dagger_{\vec{\mathbf{x}}_1} \hat{a}^\dagger_{\vec{\mathbf{x}}_2} \ldots f(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \ldots, \vec{\mathbf{x}}_a, \vec{\mathbf{x}}_b, \ldots) \hat{a}_{\vec{\mathbf{x}}_a} \hat{a}_{\vec{\mathbf{x}}_b} \ldots + \ldots . \tag{659.1}$$

So we start with particles at $\vec{\mathbf{x}}_a$ and $\vec{\mathbf{x}}_b$ and so on and finish with particles at $\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2$ and so on. The time evolution operator $\hat{\mathcal{L}}$ contains terms which destroy that initial state of particles and creates the final state. The function f connects the input and output states. We don't currently know what it is, but if the universe is to be regular and rational we do know that there must be something there. In principle, there are an infinite number of such terms, so this expression by itself is not very useful. Can we narrow it down a bit?

CONJECTURE 15.6.1. *The object $\hat{\mathcal{L}}$, which describes how God evolves the universe, can be expressed in only a finite number of ways.*

Proof:
Note that $\hat{\mathcal{L}}$ shows how God evolves the universe. It acts on one (unspecified) state vector and generates another state vector which only depends on the original vector and God. Since the same $\hat{\mathcal{L}}$ is used whatever the input vector, it cannot depend on the state of the universe. $\hat{\mathcal{L}}$ only depends on God, or, more precisely, God's relationship with matter. Therefore, it inherits the properties of timelessness, omnipresence, etc. from God. And this restricts what it can be. Since God is timeless, it doesn't matter when in time a particle is located (all other things being equal); it must relate to God in the same way and therefore its evolution must be guided by the same operator $\hat{\mathcal{L}}$. Therefore $\hat{\mathcal{L}}$ cannot (for example) depend on an absolute time scale, only relative differences in time. We can repeat this argument for the numerous different properties of God.

For this section, I will assume that there are three space dimensions and one time dimension, with a hyperbolic geometry. In principle, there is no reason why God should be restricted to this, but equally there is no reason why he should not choose it. This is one of the places where God is free to make a choice while setting up the universe.

The mathematical structure required that $e^{\frac{i}{\hbar}\hat{\mathcal{L}}\delta t}$ should be a unitary operator. This in itself restricts $\hat{\mathcal{L}}$ to be Hermitian each term must contain the same number of creation and annihilation operators, or there must be a conjugate term added to it with creation and annihilation operators swapped.

Firstly, when mapping space to a geometry, we introduce various additional degrees of freedom. These are contained within the geometrical representation, but they are something artificial we have introduced to help us understand the situation. They are not within the mind of God. The first of these is a length scale. We have labelled our particles according to some coordinate system $\vec{\mathbf{x}}_1$, $\vec{\mathbf{x}}_2$. These numbers represent the distance from some origin. But we could just as easily used a different distance scale, and have multiplied all these quantities by a factor of (say) 2. Each x is therefore implicitly multiplied by some unit of distance (let's call it $[x]$), but reality should not depend on any distance. The same thing applies for units of time, and I will use the same unit $[x]$ to measure time (this might be time multiplied by some constant speed). Now, $\langle 0 | \int d^3x \hat{a}^\dagger_{\vec{\mathbf{x}}} \hat{a}_{\vec{\mathbf{x}}} | 0 \rangle$ represents a probability, which much be dimensionless, so $\hat{a}$ and $\hat{a}^\dagger$ must have units of $[x]^{-3/2}$. This is slightly different for a Boson field; in practice here the probability density is $\int d^3x \hat{\phi}^\dagger \partial_t \hat{\phi}$ for a Bosonic creation operator $\hat{\phi}$, meaning that the Boson field has dimensions $[x]^{-1}$. Now the choices of God in

moving from one set-up of the universe to another are summarised by the quantity $\hat{\mathcal{L}}\delta t$, and this should not depend on the scale. This is a little bit naive: one scale dependent quantity is possible, since without anything to compare it against, everything else is meaningless. But every piece of scale dependence has to ultimately enter in with this one term. I'm going to save this freedom up until later.

The second artificial quantity we have introduced is the origin of the coordinate system, both in terms of space and of time. God is eternal and infinite: He exists outside both space and time, which means that for God there can be no origin. So $\hat{\mathcal{L}}$ cannot depend on where we place this origin. This means that it can only be a function of the relative separations between particles, since an absolute position is coordinate dependent. Furthermore, $\hat{\mathcal{L}}\delta t$ must be the same for each location in space and time. Since God is outside space and time, He can't (under normal circumstances) give preference to one place ahead of another. So if there is a term dependent on creation operators at a location $\vec{\mathbf{x}}_1$ in $\hat{\mathcal{L}}\delta t$, the same term must be present for each other place in space and time. So $\hat{\mathcal{L}}$ should be structured as a sum over all points in space; the equivalent of a sum for a continuous variable is an integral, which means that $\hat{\mathcal{L}}$ must be of the form,

$$\int \hat{\mathcal{L}}dt = \int d^3xdt(\ldots + \hat{a}^\dagger_{\vec{\mathbf{x}}}\hat{a}^\dagger_{\vec{\mathbf{x}}+\vec{\mathbf{x}}_2}\ldots f(\vec{\mathbf{x}}_2,\ldots,\vec{\mathbf{x}}_a,\vec{\mathbf{x}}_b,\ldots)\hat{a}_{\vec{\mathbf{x}}+\vec{\mathbf{x}}_a}\hat{a}_{\vec{\mathbf{x}}+\vec{\mathbf{x}}_b}\ldots + \ldots). \tag{660.1}$$

However, this is not quite right, because it makes an additional assumption which we are not equipped to make: namely that the physical states are localised in one place. There will be some set of states $|\psi\rangle$ which correspond to physical particles, and these will have creation operators, but they need not be localised in one place. Indeed, they cannot be: our knowledge of reality is not precise enough to say that the states are located precisely, and in practice cannot be that precise (from the uncertainty principle, which I do not intend to prove in this section, but which in practice can be derived from the same principles). I will label the creation operator for ψ as $\hat{\psi}^\dagger$, and this will be a linear superposition of the single location creation operators: $\hat{\psi}^\dagger = \hat{a}^\dagger_{\vec{\mathbf{x}}}\langle\vec{\mathbf{x}}|\psi\rangle$, where $\langle\vec{\mathbf{x}}|$ represents a particle located at position $\vec{\mathbf{x}}$. Again, there is nothing in the notation that stops us from making ψ localised in space (a point like particle); the notation is just generalised so as to not force us to make ψ localised. This $\hat{\psi}^\dagger$ will be a function of location. I will also later need $\hat{\bar{\psi}} = \hat{\psi}^\dagger\hat{\beta}$. As in my discussion of field theory in chapter 10, it is convenient to redefine the operators so that function $\hat{\beta}$ is taken out of the commutation relations and into the definition of the operator. We need that $\hat{\beta}$ somewhere, but putting it into the operator is more convenient. Once again, these are Fermions, so $\hat{\psi}\hat{\psi} = 0$.

Now suppose that the creation and annihilation operators are continuous functions of $\vec{\mathbf{x}}$. This should be the case: if they were discontinuous, then the location of that discontinuity would be a special point in space and time, and that is forbidden to us. Let us suppose that we have a continuous function $f(x)$, which we resolve at $x_1 + x_2$. We know that

$$f(x_1 + x_2) = f(x_1) + x_2\frac{df(x_1)}{dx} + \frac{x_2^2}{2}\frac{d^2f(x_1)}{dx^2} + \ldots, \tag{660.2}$$

with every additional term a higher derivative of f with respect to x. This means that since $\hat{\mathcal{L}}$ is only a function of relative location (i.e. x_2), it can be expressed in terms of derivatives of $\hat{\psi}^\dagger$ possibly multiplied by some length. Scale invariance, however, means that $\int \hat{\mathcal{L}}dt$ must not depend on the units $[x]$. This reduces the possibilities still further. I will introduce a length scale Λ, which has dimensions $[x]^{-1}$. We are now restricted to,

$$\int \hat{\mathcal{L}}dt = \int d^3xdt\hat{\psi}^\dagger(\lambda^{(0)}\Lambda + i\lambda^{(1)}_\mu\partial_\mu)\hat{\psi} + \Lambda\int d^3xdt\hat{\psi}^\dagger(\lambda^{(2)}_{\mu\nu}\partial_\mu\partial_\nu)\hat{\psi} + \Lambda^2\int d^3xdt\hat{\psi}^\dagger(\lambda^{(3)}_{\mu\nu\rho}i\partial_\mu\partial_\nu\partial_\rho)\hat{\psi} + \ldots + \Lambda^2\int d^3xdt\hat{\psi}^\dagger\hat{\psi}^\dagger\hat{\psi}\hat{\psi} + \ldots. \tag{660.3}$$

I have added in the final term with multiple creation operators to illustrate a point: that it should not be there. We know from the exclusion principle that $\hat{\psi}^\dagger\hat{\psi}^\dagger = 0$. Therefore the only terms contributing to $\hat{\mathcal{L}}$ have one $\hat{\psi}$ and one $\hat{\psi}^\dagger$. Is this right? is it true that $\hat{\psi}^\dagger\partial_\mu\hat{\psi}^\dagger = 0$? Yes, because we have the freedom to change the basis (conjecture 15.3.7). We can, for example, rotate to a basis where the creation operators are eigenvectors of the differential operator; let's call the eigenvalue distinguishing these eigenstates $\vec{\mathbf{p}}$. In this case $\int d^3x\ \hat{a}^\dagger_{\vec{\mathbf{x}}}\partial_\mu\hat{a}^\dagger_{\vec{\mathbf{x}}} = \int \frac{d^3p}{(2\pi)^3}\ \hat{a}^\dagger_{\vec{\mathbf{p}}}p_\mu\hat{a}^\dagger_{\vec{\mathbf{p}}} = 0$. The factor of $\frac{1}{(2\pi)^3}$ arises from the mathematical details I have omitted. This is now directly the product of two Fermion creation operators, so gives zero.

Going back to equation (658.3), I have introduced various dimensionless constants which I have labelled $\lambda^{(1)}_\mu$ and so on. Note that there can only be one length scale in the equation: if there were more than one, we could just treat it as some constant times the scale Λ we have introduced. However, physics itself cannot explicitly be a function of Λ. Somehow every calculation we perform has got to have Λ cancel out of it. The easiest way of doing this is to reduce the expression to just the term which contains no length scale except just a single point term which I will label with m,

$$\int \hat{\mathcal{L}}dt = \int d^3xdt\hat{\psi}^\dagger(i\lambda^{(1)}_\mu\partial^\mu + m)\hat{\psi}, \tag{661.1}$$

however we will come back to this a little bit later.

The next thing we note is that there is no preferred direction, and also no preferred velocity. God relates to each particle the same no matter which direction they are facing in, or what velocity they are travelling in: this is another consequence of God's existence outside time and space. No preferred direction implies that the system should be symmetric under rotational symmetries; no preferred velocity increases this symmetry to the Lorentz group. This initially looks a little bit difficult, since the constant $\lambda^{(1)}_\mu$ implies a direction vector, and if we get rid of that we have removed every term in $\hat{\mathcal{L}}$. For example, suppose we rotate the coordinates in equation (659.1), so that $\partial^\mu \to R^\mu{}_\nu\partial^\nu$ where R is a rotation matrix:

$$\int \hat{\mathcal{L}}dt \to \int d^3xdt\hat{\psi}^\dagger(i\lambda^{(1)}_\mu R^\mu{}_\nu\partial^\nu + m)\hat{\psi}, \tag{661.2}$$

This expression seems to depend on the rotation matrix, which we have forbidden. The constants $\lambda^{(1)}_\mu$ should be just that - constants and not vary according to our coordinate system, so we can't just rotate that direction to compensate.

However, there is a way around this, which is to increase the number of degrees of freedom for the creation and annihilation operators. For this to work, we need to have $\hat{\psi}^\dagger$ to represent at least 4 types of particles, so the simplest case leaves them as 4-component spinors rather than just simple operators. Two of these degrees of freedom refer to particles, and the other two anti-particles. So we add to $\hat{\psi}^\dagger$ an index to indicate this new degree of freedom, so it is now $\hat{\psi}^\dagger_s$. We can have a 4×4 matrix in place of $\lambda^{(1)}$, γ^{s_1,s_2}_μ, and we can write,

$$\int \hat{\mathcal{L}}dt = \int d^3xdt\hat{\psi}^\dagger_{s_1}\gamma_0(\gamma^{s_1,s_2}_\mu i\partial^\mu + m)\hat{\psi}_{s_2}, \tag{661.3}$$

If we now choose γ_μ correctly, and rotate $\hat{\psi}$ and $\hat{\psi}^\dagger$ in the s indices as we rotate the coordinates, we can simulate a rotation of γ_μ. Under a rotation, the Fermion fields transform as $\hat{\psi} \to U\hat{\psi}$ where U is a unitary matrix acting on the new s index, such that $U^\dagger\gamma_\mu U = \gamma_\nu R^\dagger{}^\nu{}_\mu$. A rotation matrix satisfies $R^\dagger R = 1$ (this is the formal mathematical definition of a rotation), so by adding these additional indices to the spinor field and having the fields change in this particular way under rotation, we can construct a $\hat{\mathcal{L}}$ that does not change when we rotate the coordinate system; indeed, this is the only way in which this can be achieved. The correct choice of γ matrices anti-commutes with each other, and, as

it happens, we can choose without losing any generality γ_0 to be the same $\hat{\beta}$ we introduced in the commutation relations of the $\hat{a}^\dagger$ and $\hat{a}$ operators. Why equate γ_0 with $\hat{\beta}$ rather than γ_1 or another of the spatial γ-matrices? $\mathcal{L}$ is related to the operator which time-translates physical states, which means that it should contains the differential operator $\frac{\partial}{\partial t}$ together with the appropriate creation and annihilation operators. Since $\hat{\beta}^2 = 1$, the only way we can rearrange (659.3) to contain this term is to set $\gamma_0 = \hat{\beta}$; the rest of the structure of $\mathcal{L}$ is fixed by Lorentz invariance, which arises from the requirement that the relationship between God and any lump matter cannot change with the inertial frame of the matter, since God is outside space and time.

There is another symmetry of the system. I mentioned that if we rotated the coefficients of the state vectors by a complex phase, it makes no difference to the physics. So we can rotate $\hat{\psi}_s \to e^{i\theta(x)}\hat{\psi}_s$. Each of these creation operators at a different position corresponds to a different particle which God could create or destroy, and so each can be rotated independently of the others. This means that θ will be a real function of x. Again, the notation allows us to consider the option that θ is constant in space, but is more general than that. We can switch to the constant θ theory by taking the limit as the coupling α (defined below) tends towards zero; there is just no a-priori reason why we should restrict ourselves to such theories, so we leave open the possibility that θ could vary across space. This rotation is equivalent to performing this complex phase rotation of the state vectors. Once again, God is outside the universe, so just as God has no preferred spatial direction or location, so God should have no preferred phase – we can regard the complex phase as another dimension, albeit one which is compact and has a circular topology. Another way of expressing this is to say that each particle, no matter what its phase, direction or position is, should experience God in the same way as any other particle in the universe, simply because God does not exist within space, time, or any complex phase space that might distinguish one particle from another. Thus $\hat{\mathcal{L}}$, which describes how God interacts with the universe, should not depend on our choice of this complex phase. An expression such as $\hat{\psi}_s^\dagger\hat{\psi}_s$ is independent of the phase, but the $\hat{\mathcal{L}}$ we have above is not. Under a phase transformation, we find

$$\int \hat{\mathcal{L}}dt \to \int d^3x dt \hat{\psi}_{s_1}^\dagger \gamma_0(\gamma_\mu^{s_1,s_2} i\partial^\mu - \gamma_\mu \partial^\mu\theta + m)\hat{\psi}_{s_2}, \tag{662.1}$$

which clearly is a function of θ. Naively this symmetry, called local gauge symmetry, seems to forbid the one vaguely interesting term we have remaining in $\hat{\mathcal{L}}$ (the differential), which would lead to a static universe. This, of course, violates God's purpose in creating a universe with rational creatures; since without any interaction or change in matter (which is what only having the m term in this $\hat{\mathcal{L}}$ implies; God would only be able to create what He had just annihilated) nothing could be said to exist, and certainly not the rational creatures necessary if God seeks to maximise goodness.

But there is another trick we can use, which is to introduce another field into the equation, which transforms simultaneously with the creation operators in such a way that cancels out the θ dependence. We will introduce a creation operator for this field $\hat{A}^\dagger$ and an annihilation operator $\hat{A}$; to maintain Hermiticity of $\hat{\mathcal{L}}$ we need to add both of these to the function, and we get

$$\int \hat{\mathcal{L}}dt = \int d^3x dt \hat{\psi}_{s_1}^\dagger \gamma_0(\gamma_\mu^{s_1,s_2}(i\partial^\mu + \alpha\hat{A}_\mu + \alpha\hat{A}_\mu^\dagger + m)\hat{\psi}_{s_2}. \tag{662.2}$$

α is some dimensionless constant, which is arbitrary. Now, if after a θ transformation, $\hat{A}_\mu + \hat{A}_\mu^\dagger \to \hat{A}_\mu + \hat{A}_\mu^\dagger + \frac{1}{\alpha}\partial_\mu\theta$, we maintain something which is invariant under gauge symmetry. $\hat{A}_\mu$ is a Boson field, and has dimensions $[x]^{-1}$. I will call it a gauge field.

However, having introduced this new field, there are another two terms we can have which are consistent with the symmetries, which only depend on this $\hat{A}_\mu$ field. We define,

$$A_\mu = \hat{A}_\mu + \hat{A}_\mu^\dagger$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{663.1}$$

and $\epsilon_{\mu\nu\rho\sigma}$ as the totally anti-symmetric tensor. If two or more of the indices for ϵ are the same, then it is 0, e.g. $\epsilon_{1123} = 0$; $\epsilon_{0123} = 1$, and every time we swap two indices we multiply ϵ by -1. So, swapping the first two indices, we find $\epsilon_{1023} = -1$, and swapping the last two indices from this quantity we find $\epsilon_{1032} = 1$, and so on. Then our full expression for $\hat{\mathcal{L}}$ is,

$$\int \hat{\mathcal{L}} dt = \int d^3x dt \left[\theta_1 F_{\mu\nu} F^{\mu\nu} + \frac{\theta_2}{2} \epsilon_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma} + \hat{\bar{\psi}}_{\vec{\mathbf{x}},s_1} (\gamma_\mu^{s_1,s_2} (i\partial^\mu + \alpha A_\mu) \hat{\psi}_{\vec{\mathbf{x}},s_2} \right], \tag{663.2}$$

where θ_1 and θ_2 are another two dimensionless constants, and $\hat{\bar{\psi}} = \hat{\psi}^\dagger \gamma_0$. Given our assumptions[4] this is the only possible form for $\hat{\mathcal{L}}$. ■

Comment:
I ought to make a clarification at this point. It is a generally accepted truism that two accurate observers will both see the same event correctly, regardless of (for example) their inertial frame, i.e. their relative velocity. This we might call a point of view invariance of physics. This is, however, different from what I am claiming, which is to say that the laws of physics themselves take the same mathematical form regardless of which inertial frame we are in. Take, for example, a classical mechanics theory based on the principle of least action. We have a Lagrangian, $\mathcal{L}(\vec{\mathbf{x}}, \frac{d\vec{\mathbf{x}}}{dt})$, written in terms of some coordinate system $(\vec{\mathbf{x}}, t)$, and the laws of physics are determined by the path that minimises the action, $\int dt \mathcal{L}(\vec{\mathbf{x}}, \frac{d\vec{\mathbf{x}}}{dt})$. For example, in Newtonian gravity, we might have in a spherically symmetric coordinate system for a particle of mass m, $\mathcal{L}(x, \frac{dx}{dt}) = \frac{1}{2} m \frac{d\vec{\mathbf{x}}}{dt} \cdot \frac{d\vec{\mathbf{x}}}{dt} + \frac{GMm}{\sqrt{\vec{\mathbf{x}} \cdot \vec{\mathbf{x}}}}$. Minimising the associated action gives us the usual set of solutions, one of which is that the particle moves in an elliptical orbit with the centre of the coordinate system as one of its focuses.

Now let us suppose that we shift to a different coordinate system. For example, we may rotate the axes, $\vec{\mathbf{x}}' = U\vec{\mathbf{x}}$, with U an orthogonal matrix independent of time (its inverse is its transpose). In this case, the mathematical form of the Lagrangian remains the same as before: $\mathcal{L}(\vec{\mathbf{x}}', \frac{d\vec{\mathbf{x}}'}{dt}) = \frac{1}{2} m \frac{d\vec{\mathbf{x}}'}{dt} \cdot \frac{d\vec{\mathbf{x}}'}{dt} + \frac{GMm}{\sqrt{\vec{\mathbf{x}}' \cdot \vec{\mathbf{x}}'}}$. We can solve this equation, get the same mathematical form of the solutions, and the only difference would be in the initial conditions we use when solving the equation of motion. We can say that this is a symmetry of the physical theory; the action is left unchanged by the coordinate transformation (which is not quite the same as saying that the Lagrangian is unchanged: it is possible to change the Lagrangian, while keeping the action constant; but for our present purposes we can ignore this distinction).

But suppose we use a different transformation of the coordinates, say: $\vec{\mathbf{x}}''(t) = \vec{\mathbf{x}}(t) + \vec{\alpha}(t)$, for some known function $\vec{\alpha}$ of time. Now the form of the Lagrangian is completely different. We can't just use our old solution and input different initial conditions; we have to work out the equations of motion again and solve them again. We would find (in general) that in this new coordinate system planets no longer move in ellipses with the centre of Gravity as one of the focuses. But that is fine. Exactly the same physics is described. We can always convert our final solution from one coordinate system to another. Point of view invariance only requires that each observer in each reference frame can construct some laws of physics which accurately describes their observations of the same events. It does not require that the Lagrangian is the same in each coordinate system (and we can treat the phase θ as though it were an additional compact dimension). It merely requires that,

[4] I want to highlight two of those assumptions which go beyond the original theological premises. Firstly, the simplification around equation (659.1) which was used to eliminate Λ from $\hat{\mathcal{L}}$; setting $\Lambda = 0$ might not be the only way of eliminating it from physical observables: I have not proved that. Secondly, I assumed that the Fermions are described by four component complex spinors in a $3+1$ dimensional hyperbolic geometry, and in particular there are no other changes of basis analogous to the local phase rotations which leave the probability unchanged. If either of these assumptions are violated, then we would have to modify or extend the theory.

knowing the differences in coordinate systems, we can convert from one formulation of the laws of physics to another.

Of course, classical (i.e. non-quantum) general relativity is the most sophisticated way of looking at this. Einstein's field equation itself, which links the stress-energy tensor with the curvature of space-time, does not depend on the reference frame. However, every component within the equation does. The procedure is that we write down the stress-energy tensor in our favourite coordinate system, solve it to find the space-time metric, and then from the metric we can calculate the various paths objects will transverse unless they are acted on by an external non-gravitational force. If we change the coordinate system, then almost always we will change the stress energy tensor. This is no problem in principle; we can plug this new tensor into the field equation, solve for the space time metric, and calculate the various trajectories of matter. The mathematical form of the final result in terms of our new coordinate system would be very different. But this doesn't matter, because it still describes the same thing, just in a different way. That all observers see the same physics does not mean that the laws of physics (for example, written in terms of a Lagrangian or equation of motion) must be the same for all observers.[5]

So let us consider an account of the laws of physics similar to that provided by Hume, where physics is merely the description of observed regularities. Here there is no reason for the equations of motion constructed by two different observers in two different coordinate systems should be the same. This I will call an observer-centric conception of the laws of physics. There is, then, no obvious reason in this account why the Lagrangian should have various symmetries (which would mean that different observers see the same mathematical form of the Lagrangian). There is no reason why different observers should write down the same Lagrangian just because they conduct experiments on the same events. There is nothing inconsistent with this account; it is just not what we observe in nature where the Lagrangian does have various symmetries. Of course, a symmetric Lagrangian is not ruled out by this conception of the laws of physics; but it isn't mandated by it.

Now let us consider the sort of procedure we are conducting, where the laws of physics are a description of God's upholding of the universe, but postulating a God who changes in time rather than the immutable God of classical theism. In this case, how people perceive God will depend on how they perceive time, which in turn depends on the coordinate system they use. Now, how God upholds the universe depends solely on God, so how God is seen to be upholding the universe depends solely on how God is seen. So the operator $\int d^4x\hat{\mathcal{L}}$, which describes how God is seen to be acting, will depend on how God is seen, which will depend on which coordinate system we are using. There is then no reason to say that the mathematical form of $\int d^4x\hat{\mathcal{L}}$ will be independent of our inertial frame, and reasons to suggest that it would not be. $\int d^4x\hat{\mathcal{L}}$ is (part of) how an observer perceives God in their own inertial frame; how that observer perceives God will depend on that inertial frame; therefore $\int d^4x\hat{\mathcal{L}}$ will vary from one frame to another. There will be no reason to enforce these symmetries onto physics.

But that is not what we have in classical theism. We have a God of pure actuality; immutable, unchangeable, transcending time and space, and none of this is in any way dependent on our frame of reference. No matter what coordinate system we use, our perception of God would be identical and our relationship with God would be identical. This means that $\int d^4x\hat{\mathcal{L}}$ will be the same in any coordinate system, since this is just a mathematical representation of our perception of God's general providence. We need to impose symmetries on $\hat{\mathcal{L}}$, not mere point of view invariance. It also means that we don't just need to consider global symmetries, where we translate the whole coordinate system by a fixed amount, but

[5]This is not philosophical relativism, which states that anything goes for any observer, and there is no reason why what one observer sees should affect what another sees. On the contrary, in this formulation, each observer's laws of physics may be different but they are not independent. Instead they are equivalent, since they can be converted to each other via a one to one coordinate mapping.

also local symmetries, where we perform different translations in different locations. Each local symmetry states that $\hat{\mathcal{L}}$ is unchanged by a transformation to a non-Cartesian coordinate system (or, in the case of gauge symmetry, a representation of the phases which varies in space); but there is nothing in the nature of God that implies that those in non-Cartesian coordinate systems will perceive Him differently from those who use Cartesian coordinate systems (or other non-Cartesian ones). On the contrary, the same considerations apply, and $\hat{\mathcal{L}}$ should be invariant under local as well as global symmetries.

It is not a trivial statement that the laws of physics should have these symmetries. It basically requires that the metaphysical origins of the laws, whatever that is, exists outside space and time and is immutable and unchangeable. We perhaps think that this is natural; but only because modern physics began in a Christian and thus classical theistic context. There was a change of perception which culminated in the seventeenth and eighteenth centuries from a divine upholder of physics to a divine lawgiver, with the divine understood according to classical theism, and it was natural that the laws should inherit some properties of their divine originator. Today, however, many people want to remove the divine reference point. But to do so removes the original *a priori* justification of thinking that the laws of physics originated outside space and time (and, also everything else). The only remaining justification for this idea that I am aware of which an atheist or agnostic can invoke is *a posteriori*, by noting that physics does indeed have these properties. Unless some such justification can be thought of, this means that atheism has much less predictive power than theism. □

It is worth pausing here to evaluate how far we have come so far. We have found a prescription to compute the probability of the universe evolving from one state to another given an initial and final state for the universe, God's free will, omnipotence, the existence of matter depending on God's continual upholding of it, and God's existence outside time, space and anything else which differentiates one state of matter from another. Our expression is,

$$P(\psi_2(t_2) \setminus \psi_1(t_1)) = A(\psi_2(t_2) \setminus \psi_1(t_1))A(\psi_2(t_2) \setminus \psi_1(t_1))^\dagger, \tag{665.1}$$

with

$$A(\psi_2(t_2) \setminus \psi_1(t_1)) = \frac{\langle\psi(t_0)| \sum_{\text{Path}} e^{-\frac{i}{\hbar}\int_{t_0}^{t_1} dt\hat{\mathcal{L}}} |\psi(t_1)\rangle \langle\psi(t_1)| \sum_{\text{Path}} e^{-\frac{i}{\hbar}\int_{t_1}^{t_2} dt\hat{\mathcal{L}}} |\psi(t_2)\rangle}{\langle\psi(t_0)| \sum_{\text{Path}} e^{-\frac{i}{\hbar}\int_{t_0}^{t_\infty} dt\hat{\mathcal{L}}} |\psi(t_\infty)\rangle}$$
$$\langle\psi(t_2)| \sum_{\text{Path}} e^{-\frac{i}{\hbar}\int_{t_2}^{t_\infty} dt\hat{\mathcal{L}}} |\psi(t_\infty)\rangle \tag{665.2}$$

We can use the time ordering operator T, which rearranges all the fields in to order from past to future, to write this as

$$A(\psi_2(t_2) \setminus \psi_1(t_1)) = \frac{\langle\psi(t_0)| T[\sum_{\text{Path}} e^{-\frac{i}{\hbar}\int_{t_0}^{t_\infty} dt\hat{\mathcal{L}}} \hat{\bar{\psi}}(t_1)\hat{\psi}(t_2)] |\psi(t_\infty)\rangle}{\langle\psi(t_0)| \sum_{\text{Path}} e^{-\frac{i}{\hbar}\int_{t_0}^{t_\infty} dt\hat{\mathcal{L}}} |\psi(t_\infty)\rangle}, \tag{665.3}$$

where $\hat{\bar{\psi}}(t_1)$ represents the creation operator that sets up our initial state, and $\hat{\psi}(t_2)$ the annihilation operator that checks for our desired final state. The sum over the paths needs to be expressed in terms of the creation and annihilation operators.

CONJECTURE 15.6.2. *The probability amplitude for a final state given a particular initial state can be expressed by the same diagrammatic expansion described in chapter 10.*

Proof:
The path obviously includes variations in both the Fermion and gauge fields, and I will discuss how we can incorporate these into our picture in turn. The Fermion fields we have already partly considered; earlier we wrote the sum in terms of state vectors $|\psi_n\rangle\, e^{\frac{i}{\hbar}\hat{\mathcal{L}}dt}\, \langle\psi_m|$.

Since we are considering continuous variables rather than discrete variables, we should rewrite the sum over paths as an integral over possible operators. We will write this as $\int d[\hat{\bar{\psi}}]d[\hat{\psi}]$ and call it the Fermionic measure (or the measure of the functional integral), where there is an implicit product over all possible states. The square brackets indicate an integral over functions rather than variables, and in practice this represents the integral over every possible function. It is easier to define the path integral if the Fermion fields were described by a finite number of discrete variables; so I will discuss this case first before moving on to the continuous case. The measure, like the action and for the same reasons, needs to be unchanged under the various symmetry transformations we have been considering, and this provides it with an unique and well-defined formal mathematical definition.

A path consists by identifying various fields on neighbouring time slices. So we can say that one particular path involves state $|\psi_1\rangle$ at time t_1, $|\psi_2\rangle$ at time t_2, $|\psi_3\rangle$ at time t_3, and so on. We need the creation operators for all these states to be present for the path to be realized. This means that

$$\int d[\hat{\bar{\psi}}]\hat{\bar{\psi}} = 1 \qquad\qquad \int d[\hat{\bar{\psi}}] = 0. \tag{666.1}$$

The first of these is saying 'include this path in the sum if the field is present.' The second says 'exclude this path from the sum if the field is absent.' The same rules apply for the $\hat{\psi}$ path integral. It might seem that this introduces a problem, because $\int d[\hat{\bar{\psi}}]d[\hat{\psi}]$ includes a product over all possible states, and not just those which contribute to the actual path of the particle we are interested in. But in practice this is resolved in a neat way. The exponential of $\hat{\mathcal{L}}$ contains creation and annihilation operators for particles in every state. These represent particles which come into and out of existence from nothing, I will call them disconnected particles, and the real life particle whose path we want to consider as the actual particle. When we insert a real path of an actual particle into the amplitude, we effectively exclude that from the sea of disconnected particles. So instead of our amplitude representing an empty state plus the path of the actual particle we want to track, it represents a completely full state with every possible permutation of disconnected particles and anti-particles represented with the path of the actual particle we want to track subtracted from it. So instead of tracking the path of a particle, we are tracking the path of a hole. I stated earlier that two situations of the particle and the hole were physically equivalent: mathematically there is a symmetry linking them. That is the same here. But what about all these disconnected particles being created from nothing, unrelated to our actual particles? After all, nothing comes from nothing. The point is that this is just an artefact of the mathematical representation. The disconnected particles aren't really there. The same swarm of disconnected particles is in the denominator of the amplitude as well as the numerator. Their effect is just a multiplicative factor, which cancels out between the numerator and denominator of the amplitude A (equation (663.3), leaving us with just the actual particle we want to track.

The Boson gauge fields are somewhat easier than the Fermion fields. The operators commute rather than anti-commute with each other, just like the numbers we are used to. We can parametrise the operator with set of numbers A_μ at each time slice, and to sum over the paths we just have to integrate over those numbers. The only issue is defining the measure of that integration. We want the measure to be invariant under the same symmetries as the action, i.e. gauge invariant and so on, and in practice this means that for the example we are looking at so far, we need a straight forward integral over the A field.

So we can write the amplitude as

$$A(\psi_2 \setminus \psi_1) = \frac{\int d[\hat{\bar{\psi}}]d[\psi]d[A]T[e^{-\frac{i}{\hbar}\int d^4x(\theta_1 F_{\mu\nu}^2 + \frac{\theta_2}{2}\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} + \hat{\bar{\psi}}\gamma^\mu(i\partial_\mu + \alpha A_\mu)\hat{\psi})}\hat{\bar{\psi}}_1\hat{\psi}_2]}{\int d[\hat{\bar{\psi}}]d[\psi]d[A]T[e^{-\frac{i}{\hbar}\int d^4x(\theta_1 F_{\mu\nu}^2 + \frac{\theta_2}{2}\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} + \hat{\bar{\psi}}\gamma^\mu(i\partial_\mu + \alpha\Lambda_\mu)\hat{\psi})}]} \tag{666.2}$$

Now it is time to look at the path integral in a little more detail. In particular, we can study Gaussian integrals, that is integrals of the form (where D here represents an arbitrary number)

$$\int d[\hat{\bar{\psi}}]d[\hat{\psi}]e^{\hat{\bar{\psi}}D\hat{\psi}} = \int d[\hat{\bar{\psi}}]d[\hat{\psi}](1+\hat{\bar{\psi}}D\hat{\psi}+\frac{1}{2}(\hat{\bar{\psi}}D\hat{\psi})^2+\ldots) = D. \tag{667.1}$$

Here I just have one single ψ state. What if we have several states connected by a matrix? Here the trick is to reduce it to the problem above, by rotating the basis into the eigenvalues of the matrix. So if the eigenvectors of the matrix D are $\tilde{\psi}_i$, and eigenvalues D_k, we can write $\hat{\psi}_a = \sum_k \tilde{\psi}_k \langle\tilde{\psi}_k|\hat{\psi}_a\rangle$, where the sum is over the complete set of eigenvectors. This means that

$$\begin{aligned}
\int \prod_i d[\hat{\bar{\psi}}_i]d[\hat{\psi}_i]e^{\hat{\bar{\psi}}_i D_{ij}\hat{\psi}_j} &= \int \prod_i d[\bar{\psi}_i]d[\hat{\psi}_i]e^{\langle\hat{\bar{\psi}}_i|\bar{\tilde{\psi}}_k\rangle\bar{\tilde{\psi}}_k D_{ij}\tilde{\psi}_k\langle\tilde{\psi}_k|\hat{\bar{\psi}}_j\rangle} \\
&= \int \prod_i d[\bar{\psi}_i]d[\hat{\psi}_i]e^{\langle\hat{\bar{\psi}}_i|\bar{\tilde{\psi}}_k\rangle\bar{\tilde{\psi}}_k D_{ij}\tilde{\psi}_m\langle\tilde{\psi}_m|\hat{\bar{\psi}}_j\rangle} \\
&= \int \prod_i d[\bar{\tilde{\psi}}_i]d[\tilde{\psi}_i]e^{\bar{\tilde{\psi}}_k D_k\tilde{\psi}_k} \\
&= \int \prod_i d[\bar{\tilde{\psi}}_i]d[\tilde{\psi}_i](1+\bar{\tilde{\psi}}_1 D_1\tilde{\psi}_1+\ldots)(1+\bar{\tilde{\psi}}_2 D_2\tilde{\psi}_2+\ldots)\ldots \\
&= \prod D_k = \det(D)
\end{aligned} \tag{667.2}$$

We have used the idea that $\langle\hat{\bar{\psi}}_i|\bar{\tilde{\psi}}_k\rangle D_{ij}\langle\tilde{\psi}_m|\hat{\bar{\psi}}_j\rangle$ is a diagonal matrix, which follows from the definition of the eigenvectors. The integral over the fields is then just reduced to one integral of the form (665.1) for each eigenvector. We have one functional integral for each state in the basis, so one integral for each eigenvector. We know how to solve those integrals - the result is the product of the eigenvalues, and the determinant of a matrix denoted $\det(D)$ is defined as the product of the eigenvalues. $\det(D)$ represents the (fictitious) swarm of disconnected particles I was discussing earlier.

Now we insert some initial and final Fermion states into the integral, and we find

$$\begin{aligned}
\int \prod_i d[\hat{\bar{\psi}}_i]d[\hat{\psi}_i]\hat{\bar{\psi}}_a\hat{\psi}_b e^{\hat{\bar{\psi}}_i D_{ij}\hat{\psi}_j} &= \int \prod_i d[\hat{\bar{\psi}}_i]d[\hat{\psi}_i]\langle\bar{\tilde{\psi}}_k|\hat{\bar{\psi}}_a\rangle\bar{\tilde{\psi}}_k\langle\tilde{\psi}_m|\hat{\psi}_b\rangle\tilde{\psi}_m e^{\bar{\tilde{\psi}}_i D_{ij}\hat{\psi}_j} \\
&= \det D\sum_k \langle\bar{\tilde{\psi}}_k|\hat{\bar{\psi}}_a\rangle\frac{1}{D_k}\langle\tilde{\psi}_k|\hat{\psi}_b\rangle \\
&= \det D\frac{1}{D_{ab}}.
\end{aligned} \tag{667.3}$$

So what has happened here is that we have expressed both our initial and final states in terms of the eigenvector basis, and we have done the same trick inside the exponential. Once again, the exponential reduces to the product of terms such as $e^{\bar{\tilde{\psi}}_k D_k\tilde{\psi}_k}$ which we can expand as $1+\bar{\tilde{\psi}}_k D_k\tilde{\psi}_k+\frac{1}{2}(\bar{\tilde{\psi}}_k D_k\tilde{\psi}_k)^2+\ldots$. Every term in this expansion which contains two or more creation operators for the same eigenstate gives zero because the square of creation operators for Fermion states gives zero. Every term with no creation operators for a particular eigenstate gives zero because of how we defined the functional integral in equation (664.1). So, once again, the final result is the product of the eigenvectors D_k. But this time, we have the initial and final states to worry about as well. For one of the eigenvectors, we get the required $\tilde{\psi}_k$ and $\bar{\tilde{\psi}}_k$ operators from the initial and final states rather than from the expansion of the exponential. And when these operators appear as part of the initial and final states they are not accompanied by the eigenvalue D_k. The net result is therefore we get the product of all the eigenvectors except the one associated with the initial and final states, so $\frac{\det D}{D_k}$. Rotating back into the original basis then leads to $\det D\frac{1}{D_{ab}}$,

where $\frac{1}{D_{ab}}$ represents the component of the inverse of the operator D which links the states a and b. For example, if D were a 2×2 matrix, and $a = 1$, $b = 2$, we would find the inverse of the matrix, and then find component in the first row and second column of that inverse.

$\frac{1}{D_{ab}}$ is known as the propagator to go from state a to state b, and is going to play a crucial role in a moment. When we have several Fermion fields, the procedure is similar. Once again, we have to replace each pair of $\hat{\bar{\psi}}$ and $\hat{\psi}$ Fermion states with a propagator. However, there are numerous different ways in which we can pair up the Fermion fields, and we need to sum over the possible arrangements of the propagators, being careful about minus signs from anti-commuting the fields. For example, if we are interested in a state with two particles,

$$\hat{\bar{\psi}}_a \hat{\bar{\psi}}_b \hat{\psi}_c \hat{\psi}_d,$$

Then there are two options, to pair together the a and d fields plus the b and c fields, or to pair together a and c plus b and d. We replace each pair with a propagator, giving

$$\int \prod_i d[\hat{\bar{\psi}}_i] d[\hat{\psi}_i] \hat{\bar{\psi}}_a \hat{\bar{\psi}}_b \hat{\psi}_c \hat{\psi}_d e^{\hat{\bar{\psi}}_i D_{ij} \hat{\psi}_j} = \left(\frac{1}{D_{bc}} \frac{1}{D_{ad}} - \frac{1}{D_{bd}} \frac{1}{D_{ac}}\right) \det D. \tag{668.1}$$

The minus sign in front of the second pair of fields is because we had to swap the order of the fields to contract them, and every swap of two Fermion states implies a multiplication by -1. When we compute the probability amplitude, we divide the path integral with initial and final states by the path integral without those states, which has the effect of cancelling out that factor of $\det D$. We are just left with the propagator.

For a Bosonic field, we find something similar,

$$\begin{aligned} \int d[A] e^{A_i F_{ij} A_j} &= \frac{(\sqrt{\pi})^N}{\sqrt{\det F}} \\ \int d[A] A e^{A_i F_{ij} A_j} &= 0 \\ \int d[A] A_a A_b e^{A_i F_{ij} A_j} &= \frac{(\sqrt{\pi})^N}{\sqrt{\det F}} \frac{1}{F_{ab}}, \end{aligned} \tag{668.2}$$

where N represents the number of possible Boson states. Once again the determinant is cancelled out by the denominator of the amplitude. The procedure when there are several fields is the same as in Fermions, except that there are no minus signs.

$$\int d[A] A_a A_b A_c A_d e^{A_i F_{ij} A_j} = \frac{(\sqrt{\pi})^N}{\sqrt{\det F}} \left(\frac{1}{F_{ab}} \frac{1}{F_{cd}} + \frac{1}{F_{ac}} \frac{1}{F_{bd}} + \frac{1}{F_{ad}} \frac{1}{F_{bc}}\right). \tag{668.3}$$

All the terms we have considered so far with one creation operator and one annihilation operator contribute to $\hat{\mathcal{L}}$ (up to various numerical factors), but there is one more term which has has a Fermion creation operator, Fermion annihilation operator, and the gauge Boson operator: $\hat{\bar{\psi}} \alpha \gamma_\mu A^\mu \hat{\psi}$. We can treat this in a similar way to the Fermion terms. When we expand the exponential of the part of $\hat{\mathcal{L}}$ containing the Fermions, in addition to the contributions giving us the propagator as before, we also have a series involving various powers of $\alpha \hat{\bar{\psi}} \gamma_\mu A^\mu \hat{\psi}$. Each of these terms represents the destruction of a Fermion field, creation of a second Fermion field, and either destruction or creation of a gauge field (which we will call a photon); or alternatively a photon field decaying into a Fermion and anti-Fermion field, or a Fermion and anti-Fermion field decaying into a photon. The Fermion creation and annihilation operators will lead to two end points of a quark propagator, which are matched with one beginning or end of a photon propagator. We then perform the integral (666.2) over the gauge field, and these photon operators will perform the role of initial and final photon states. For example we can propagate from our initial state to a place where we create a photon, propagate the photon and Fermion to a place where the photon is annihilated, and then propagate the Fermion to our final state. We need to

sum over every possible event where a photon can be created or annihilated, and each of these will contribute to the amplitude. The finer details of the calculation (which I have omitted) show that momentum – represented by the eigenvalue $\vec{\mathbf{k}}$ is conserved at each of these creation/annihilation events. This is a consequence of the locality of $\mathcal{L}$ that it only depends on relative separations and thus differential operators. This means, of course, that we have recovered the laws of efficient and final causality - each particle can only decay into certain other particles, and has a particular sequence of particles leading up to it. Nothing can come from nothing. Thus it is fully consistent to treat the particles as secondary causes, and in this way we avoid occasionalism which denies that this is possible.

What we have done here (I have omitted numerous details – the polarisation of light and necessity of gauge fixing being the most important) is ended up in exactly the same place as described in chapter 10. The subsequent diagrammatic expansion and amplitudes we have calculated are exactly the same as for unrenormalised quantum electrodynamics as described in that chapter: the starting point might have been different, but the final result is identical, down to the smallest detail. ■

Discussion:
But as we saw in chapter 10, the unrenormalised theory is not much good - it gives infinite probabilities and other nonsense such as that. So what has gone wrong?

Early in my argument (equation (659.1)) I invoked scale invariance as a key principle, and stated that the easiest way of achieving this was to construct the Lagrangian so that it was dimensionless, which left only one possible term, the $\int d^4x\hat{\psi}^\dagger\gamma_0\gamma^\mu(i\partial_\mu + \alpha A_\mu)\hat{\psi}$ term I have been using. But there is a better way of both understanding and enforcing scale invariance. This requires our freedom to rotate the Fermion fields to another basis (conjecture 15.3.7). As long as the measure of the functional integration remains invariant under gauge, translation and Lorentz symmetries, we can freely rotate the Fermion fields by some operator U, so $\hat{\psi} \to U\hat{\psi}$. If we make U a function of the derivative operator $\frac{1}{\Lambda}(i\partial_\mu + \alpha A_\mu)$, then this transformation changes the Lagrangian. The new Lagrangian will contain every possible term consistent with Lorentz and gauge symmetry:

$$\hat{\mathcal{L}} \to \lambda_1^{(1)}\hat{\bar{\psi}}\gamma^\mu(i\partial_\mu + A_\mu)\hat{\psi} + \frac{\lambda_2^{(1)}}{\Lambda^{(1)}}\hat{\bar{\psi}}\gamma^\mu(i\partial_\mu + A_\mu)\gamma^\nu(i\partial_\nu + A_\nu)\hat{\psi} + \frac{\lambda_3^{(1)}}{\Lambda^{(1)}\Lambda^{(1)}}\hat{\bar{\psi}}\gamma_\mu(i\partial_\mu + A_\mu)\gamma^\nu(i\partial_\nu + A_\nu)\gamma^\rho(i\partial_\rho + A_\rho)\hat{\psi} + \dots . \quad (669.1)$$

So by doing this, we have turned a simple Lagrangian – just one term – into a complete mess, and in particular introduced an infinite number of new constants into our theory, the dimensionless $\lambda^{(1)}$s and the scale $\Lambda^{(1)}$ (in practice, one of the λs is redundant, so we can set one of them to be 1). Now, if we apply the transformation on the Fermion fields again, what will happen is that we change the Lagrangian again: we will change the values of the parameters to a different set $\lambda^{(2)}$ and the scale to $\Lambda^{(2)}$. And if we apply the transformation again, we will once again change all the parameters. You might think that each of these different Lagrangians should be physically equivalent rotating the Fermion basis doesn't change any of the physics. However, it is not quite the same: we have rotated the Lagrangian, and the measure, but not the operators representing our initial and final states. Instead of the particles created by $\hat{\psi}$, we can now only easily compute amplitudes and probabilities for the rotated fields $U\hat{\psi}$. So that is what this rotation does: if the physical particles are not created by the simple operators we started with, then we need to perform this rotation of basis until the creation operators in the Lagrangian $\hat{\mathcal{L}}$ match up with the creation operators corresponding to the physical particles we see in our experiments.

Now let us consider an infinitesimal transformation of the fields, which we apply repeatedly. The sequence $\lambda_1^{(1)} \to \lambda_1^{(2)} \to \lambda_1^{(3)}$ is broken down into a continuous transformation of these parameters. We introduce a new parameter τ which counts how many times we have

applied the transformation, and we can then write the parameters as continuous functions of τ: $\lambda_1^{(\tau)}$. As we apply the transformation, these parameters will evolve smoothly according to a differential equation that can be computed from knowledge of the rotation U. So we have an infinite number of parameters, each subject to a differential equation describing their evolution in τ. This is mathematically the same situation we have in classical mechanics, τ being the analogue of time and the λ parameters and the scale Λ representing the positions of various particles. But if this process is just an analogue of a complicated classical mechanics system, then the second law of thermodynamics still applies: choose the right transformation, apply it enough times, and eventually you reach an equilibrium state where all the parameters remain the same after the transformation. There is one equilibrium state where the parameters in the Lagrangian are independent of this transformation of the fields.

Now each application of the transformation changes the scale Λ. This Λ has the dimensions of inverse length, and represents some fundamental length scale in our theory, called the correlation length. Our transformation either shrinks or expands that length scale. The physical interpretation of the transformation U is that it symbolically represents a change in this purported fundamental physical scale. But there is an equilibrium set of parameters – let us call them $\lambda^{(*)}$ and the Lagrangian $\hat{\mathcal{L}}_*$ – which do not change under the transformation. Our initial premise was that physical observables must be blind to any such length scale because God is outside of time and space and therefore to Him every length or duration is the same as any other: there are only relative, not absolute, length scales. The set of parameters $\lambda^{(*)}$ describe a theory where the physical particles don't care about changes in the correlation length – their relationship with God, described by the Lagrangian and in particular its coefficients – remains the same no matter what the correlation length actually is. But that is exactly the theory which we want. The dimensionless Lagrangian we have been using is not scale invariant in this sense because it is not invariant under the scale changing transformation U. But $\hat{\mathcal{L}}_*$ is, and that's the Lagrangian which describes physics. Similarly, the physical particle basis is that one which you get after applying the scale-changing transformation a large number of times, so that it becomes a fixed point unchanged under subsequent applications of the transformation operator: the operator that represents physical particles is the one that remains unchanged under this transformation of the coordinate system. All we have to do is pick any transformation that satisfies the basic conditions – the measure of the path integral $[d\hat{\bar{\psi}}][d\hat{\psi}]$ needs to be invariant under the appropriate transformations, for example – and which represents a monotonic change in the scale, apply it an infinite number of times, and we can construct the physical Fermion states from any other basis of creation and annihilation states.

The theory described by $\hat{\mathcal{L}}_*$ allows us to compute sensible probabilities. This process of repeatedly transforming the fields is known as renormalization. The problem is that we now have an impossibly complicated Lagrangian, but the good news is that we can perform the rotations of the fields after computing the amplitudes rather than beforehand. So we modify the 'simple' Lagrangian (661.2) so that the infinities are regulated by some parameter which we can express as $1/\epsilon$, compute the probability amplitudes, rotate the fields to remove all divergent dependency on ϵ, and then take the limit as $\epsilon \to 0$. The calculations are still difficult, but they are at least possible. It can be done systematically, and it compares exceptionally favourably with experiment. □

15.7. The Classical Limit

CONJECTURE 15.7.1. *Classical mechanics is the $\hbar \to 0$ limit of this theory.*

Proof:
The picture discussed above, where particles can in principle travel by any possible path from A to B, so we have to consider all paths when constructing a probability amplitude, seems to differ from the classical mechanical picture which states that there is a single possible path determined by the principle of least action. However, classical mechanics emerges from this picture in a very neat way. The quantum field theory paths are weighted by the function $e^{\frac{i}{\hbar}\int d^4x \hat{\mathcal{L}}^{\text{Path}}}$. Now suppose that the paths are distinguished from each other by a single parameter s. This is a simplification, but the generalisation of this argument to an arbitrary number of parameters is straight-forward. Suppose that we choose one path with some particular $s = s_0$. Then the weighting factor for the neighbouring path is given by

$$e^{\frac{i}{\hbar}\int d^4x \left[\hat{\mathcal{L}}^{s_0} + \frac{\partial \hat{\mathcal{L}}^{s_0}}{\partial s}(s-s_0) + \frac{1}{2}(s-s_0)^2 \frac{\partial^2 \hat{\mathcal{L}}^{s_0}}{\partial s^2}\right]}. \tag{671.1}$$

Now if $\frac{\partial \hat{\mathcal{L}}^{s_0}}{\partial s}(s - s_0)$ is close to $\pi\hbar$, then there will be complete destructive interference between the neighbouring paths, and as $\hbar$ tends towards zero, the variation in the path needed to have this interference will become closer and closer. What this means is that when contributions of the very similar paths are added together in the amplitude, their effects will cancel out. However, there is one exception to this rule, and that is when $\frac{\partial \hat{\mathcal{L}}^{s_0}}{\partial s} = 0$. In this case, the variation in the weight between neighbouring paths is much smaller; we have to turn to the $(s - s_0)^2$ term to get a contribution. So if we choose s_0 to represent the path that minimises $\hat{\mathcal{L}}$, then for those paths where (as a rough approximation) $\frac{1}{2}(s-s_0)^2 \frac{\partial^2 \hat{\mathcal{L}}^{s_0}}{\partial s^2} < \frac{\pi\hbar}{2}$ will interfere constructively, while any paths which deviate from s_0 by a larger amount than this will come in with both positive and negative weights, and their effects will more or less cancel out. As $\hbar$ decreases, the allowed deviation from the path that minimises $\hat{\mathcal{L}}$ also decreases, and when $\hbar$ reaches zero, only the path that minimises $\hat{\mathcal{L}}$ will contribute to the path integral.

In other words, as $\hbar \to 0$, all the particles will appear to travel along the path which minimises $\int d^3x dt \hat{\mathcal{L}}$. Now all we have to do is identify $\int d^3x dt \hat{\mathcal{L}}$ with the classical action, observe that $\hbar$ is in practice very small in comparison to the sort of distance and momentum scales we are used to, and we have classical mechanics (or, at least, Lagrange's formulation of it using the principle of least action). $\hbar$ has dimensions of distance times momentum, so the sort of scale where quantum effects become important is when $\Delta x \Delta p \sim \hbar$, where Δx is the difference in location and Δp a momentum difference between two particles we are studying. In the 'Real world', the sort of distances and momenta we see around us are much larger than this sort of scale. Classical mechanics holds to a very good approximation. However, when we reduce the distance considerably further, and get down to the size of atoms, then the non-zero value of $\hbar$ becomes important. Classical mechanics breaks down, and we need to use quantum field theory. ■

Discussion:
What is particularly interesting about this is that the philosophical bases of classical and quantum mechanics are very different, yet they lead to very similar predictions, at least at 'real world' length scales. We can fly rockets to the moon, play sports, fire artillery using classical mechanics, and the fact that it is the wrong theory is something we would never notice: at these sort of scales, the difference between quantum and classical mechanics is so small that only a ridiculously sensitive experiment would notice it. But the philosophical implications of this give a lesson which we should be careful to heed. Classical mechanics is deterministic, and this lead people to believe that we needed a deterministic metaphysics. However, it turns out that we don't; a non-deterministic metaphysics can make the same experimental predictions. Quantum field theory is a more fundamental theory than classical mechanics. We can derive classical mechanics as a limit of field theory; but the process is

one way: we cannot extract quantum field theory from classical mechanics using a similar method (as shown by Bell's inequalities). Indeed, it is difficult to see how a classical system could give rise to quantum theory. Classical systems are deterministic; how we can derive something which is indeterminate from something that is necessarily determinate is beyond me. However, an indeterminate system like field theory can nonetheless lead to something that is indistinguishable from something deterministic in one particular limit. We should not be fooled by the success of classical mechanics into thinking that the philosophy derived from it is thus necessarily sound; we can derive the same results from something very different. We must turn to the most fundamental physics available to derive our philosophy.

Equally, I see no point whatsoever in using non-relativistic quantum mechanics as the basis of our philosophical system, and then trying to adapt that interpretation to field theory, because the differences between the two are significant. We should start with field theory.

It might have been argued that the mechanistic philosophy is consistent with classical mechanics but inconsistent with quantum physics, but the Aristotelian system consistent with quantum physics but not classical physics; we thus have something of an impasse where one philosophy works in the macroscopic world and the other in the microscopic world, so we should not prefer one over the other, or use the mechanistic philosophy as good enough in our day-to-day lives, just as classical mechanics is good enough to describe everyday physics. What we have just shown, however, is that this argument fails: Aristotelian philosophy describes both the quantum and classical worlds; while mechanistic philosophy is only consistent with the classical world. There is thus no reason to prefer a mechanistic philosophy just because classical physics is an exceptionally good approximation to the world at the scales which are most natural to us. The Aristotelian world-view describes the world at these scales just as well as the mechanical world-view. But Aristotelian philosophy should be preferred because it also works in the regime of quantum field theory; it is the more general approach.

Of course, were the universe fundamentally classical (for example, by exactly obeying a least action principle), then while this would not rule out the Aristotelian metaphysics, it would leave it at a considerable disadvantage in comparison to the mechanical world view: it would have to additionally explain why $\hbar$ was sufficiently small in comparison to observed energy and time scales that its effects were below our experimental resolution. However, a non-zero $\hbar$ is fatal to the mechanical world view, at least in its most straight-forward expressions.

Of course, this begs a question. What if something more fundamental than quantum field theory replaces it? Would I then have to overthrow the philosophy I have advocated here? Well, that depends on what replaces field theory. The two key aspects I need for my philosophy are indeterminacy, that matter can be created and destroyed, and locality, with the locality giving rise to the various causality relations (or perhaps the causality relations imply the locality principle, depending on your point of view); the underlying symmetries of the physics are also important, though perhaps a little less fundamental since which symmetries are needed depends on the types of matter in the universe and the precise nature of space-time. Different worlds with different geometries of space time might have different symmetries and topologies, without making such an indent on the underlying philosophy. Anything which respects these principles would be consistent with the philosophy I have advocated here. As far as we know, the three principles are satisfied. Transformation of matter from one form (e.g. photon) into another (e.g. electron/positron pair) is too well observed for its possibility to be questioned. Is the universe indeterminate? There are philosophical models of quantum field theory which try to keep determinacy, such as the Everett Many-Worlds theory and the De Broglie-Bohm interpretation, but these seem less likely to me.

The Everett interpretation, which tries to interpret the state vectors according to frequentest probability theory by stating that the universe branches off into numerous sub-universes with each quantum event, begs too many questions to be reasonable. It keeps determinacy, but at a very high cost. These sub-universes are unobservable, so this hypothesis can never be tested experimentally; and the question arises why the multiverse should act in this way. The idea that there is one, indeterminate, universe is far simpler and far more in accord with our common sense. Plus, of course, theism can be empirically tested, because it allows for the possibility of miracles, as I discuss later; while the multiverse is inconsistent with most forms of theism (at least those which involve a God able to see into and plan what is for us the future).

The de Broglie-Bohm interpretation of quantum mechanics makes use of an underlying guide wave which moves the observed particles; however this is non-local and difficult to reconcile with quantum field theory.

Finally, substance locality, through the conservation laws related to it, has been tested extensively, and non-local systems tend to involve infinite momentums appearing every now and then, which would surely have been observed. For example, Lorentz symmetry has been tested to energies that go beyond the Planck scale, where we would expect the effects of quantum gravity to be important. I would not be surprised if our current standard model were overthrown in the future, but I would be very surprised if whatever replaces it will be sufficiently different that it will require making major changes to the philosophy. Minor adjustments to the details will certainly be required, but not enough to overthrow the general principles. Of course, I could be wrong; if so, then the scholastic system would again need to be rejected, and something else take its place. But until that day, we will just have to work on the assumption that God does, after all, play dice; though the rules of His game respect substance causality and various conservation laws. □

15.8. The Standard Model of Particle Physics

CONJECTURE 15.8.1. *Other forces apart from electromagnetism are possible.*

Proof:
It is now time to resume our derivation of physics from classical theism. So far, I think, the only experimental input I have needed to derive quantum electrodynamics in addition to my theological premises is that there are three space and one time dimensions with a hyperbolic geometry, and that the matter we see around us is dominated by Fermions and Bosons related by a gauge symmetry. This was all a free choice of God (and being free we cannot derive it from more fundamental principles - but see section 15.11 below); He could have chosen otherwise and constructed a different universe, but He did not. Now we come to other places where God had a choice in His construction of the universe. I cannot say why God choose to augment electromagnetism with the strong and weak nuclear forces and the Higgs field (apart from the considerations of section 15.11), nor why He should have three families of matter. It may be that there is a more fundamental reason which will reveal itself as physics progresses, but at the moment this just seems to be a free choice of God. But what I shall attempt to do here is show that these forces are consistent with the picture I have been developing.

A first principles derivation of the full standard model is not going to be possible, because there are some freedom in the parameters and particles which could be within it. Therefore, and for the sake of brevity, my discussion below shall take more of a description rather than a proof. However, the discussion is not a free-for all. The goal is to write down

the most general possible theory consistent with the underlying symmetries and various additional assumptions.

- There are four types of Fermion fields in three families.
- There is a $SU(3) \times SU(2) \times U(1)$ gauge symmetry.
- There are 3 space and one time dimension.
- There is a solitary scalar field which couples only to the $SU(2) \times U(1)$ gauge fields.
- The left and right handed Fermion fields couple differently to the $SU(2) \times U(1)$ symmetry.

Aside from the various dimensionless parameters of the theory which cannot be constrained by symmetry requirements, if we make these assumptions about the types of fundamental matter fields present in the universe, we are led inexorably to the standard model. God could have chosen to populate the universe with different matter fields; these assumptions each represent a free choice by God and therefore not something we can prove *a priori*, but only observe *a posteriori*, and subsequently use as the basis of our model. However, what we can then do is write down the most general theory consistent with these assumptions, and that gives us the standard model of particle physics described below – there are no other possibilities. So while it might seem as though I am plucking the theory out of the air, everything I write is in practice determined by mathematical necessity building on my premises and assumptions.

The first thing we can do is to expand the Fermion fields even further. So far, we have made them a four component spinor at each site. But why stick to four? We can add an additional set of indices. So, for example, we can give the spinors two extra indices, one which can take two values, and the other three upgrading them to twenty four complex numbers for each position. We have $a^\dagger_{\vec{\mathbf{x}},s,y,w}$, where $\vec{\mathbf{x}}$ is the location, s the spinor index we have already seen, y a three dimensional index which can take values $1, 2, 3$ and w a further index which can take the values 1 and 2. The coefficients c and $c^\dagger$ which describe how the state is built up from the particles created by these operators will also get similar indices. The probability amplitude is still $P(x) = \sum_{s,y,w} c^\dagger_{\vec{\mathbf{x}},s,y,w} c_{\vec{\mathbf{x}},s,y,w}$. This means that the amplitude is invariant under a unitary rotation of the y and w indices; in the case of the y index this rotation is provided by a 3×3 unitary matrix (the type of matrix is called SU(3), and I will label the corresponding eight gauge fields C_μ), while for the w index this is a 2×2 unitary matrix (this type of matrix is called SU(2) and I will label the corresponding three gauge fields B_μ). The electromagnetic field A_μ is still present as well. In each case, we need to introduce a corresponding gauge field just as we did for electromagnetism above, and a similar gauge field action.

The second thing which we do is expand the number of Fermion fields. We divide them into two types. Firstly, the leptons, which have the w index and not the y index. The two lepton fields are electrons (corresponding to $w = 1$) and neutrinos (corresponding to $w = 2$). The second type of Fermion fields are quarks, and these have both w and y indices. We will label the $w = 1$ quark the up quark and the $w = 2$ quark the down quark. The y index is unfortunately referred to as colour, and the three states called red, green and blue. So we can have a red up quark, or green down quark, and so on. These eight particles complete one family of Fermions. There is a good reason why we are restricted to these eight particles, which I will discuss a little later.

Why just have one family of Fermions? There is no good reason to do so, so I will choose to have another two families to go with it. These families I will indicate with the index f, which can take values 1, 2 or 3, these are known as flavours. In nature, there are a couple of additional complications: the B field representing the weak nuclear force acts differently on different spinor indices (the two different ways of writing this are known, again for unfortunate historical reasons, as left and right handed fields, and are eigenstates of the operator $P_\pm = \frac{1}{2}(1 \pm \gamma_0\gamma_1\gamma_2\gamma_3)$), and the strong force (the SU(3) force) acts on the different flavours of Fermions in a slightly different way to the weak and electromagnetic

forces. All of these are allowed according to symmetries, but just make the physics a little more complicated; under our principle that we should keep things as general as possible given the nature of God and demands of logical consistency so as not to unduly restrict God we need to allow for these in our formalism. However, there is one simplification, which is that the possible $\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma}$ term in the action doesn't seem to be there in practice. Why this is the case remains one of the major outstanding problems of particle physics; I am not going to solve it here, but will meekly accept the experimental result. We label the quark fields with a q and the lepton fields with an l.

Let's start by constructing an action involving left and right handed massless Fermion fields. We want to split the action between left and right handed fields. I will write the derivative operator as $\not{D} = \gamma^\mu(i\partial_\mu + Q\alpha A_\mu + Q_W\alpha_W B_\mu)$, leaving open the possibility that the left and right handed fields and the leptons and quarks have different charges Q_W and Q (so, we will subscript the Dirac operator with either left or right and lepton and quark so we know which charges we are using). I will also put the Fermions into couplets (corresponding to the 2-component w index), E_L featuring an electron and a neutrino, and Q_L featuring an up and down quark,

$$\begin{aligned}
\hat{E}_L &= P_+ \begin{pmatrix} \hat{e} \\ \hat{\nu}_e \end{pmatrix} & \hat{\bar{E}}_L &= (\hat{e}^\dagger, \hat{\nu}_e^\dagger)\gamma_0 P_- \\
\hat{E}_R &= P_- \begin{pmatrix} \hat{e} \\ \hat{\nu}_e \end{pmatrix} & \hat{\bar{E}}_R &= (\hat{e}^\dagger, \hat{\nu}_e^\dagger)\gamma_0 P_+ \\
\hat{Q}_L &= P_+ \begin{pmatrix} \hat{u} \\ \hat{d} \end{pmatrix} & \hat{\bar{Q}}_L &= (\hat{u}^\dagger, \hat{d}^\dagger)\gamma_0 P_- \\
\hat{Q}_R &= P_- \begin{pmatrix} \hat{u} \\ \hat{d} \end{pmatrix} & \hat{\bar{Q}}_R &= (\hat{u}^\dagger, \hat{d}^\dagger)\gamma_0 P_+
\end{aligned}$$

so $\hat{e}$ represents the annihilation operator for the 4 component electron spinor, $\hat{\nu}_e$ for the electron neutrino, $\hat{u}$ for the up quark, and $\hat{d}$ for the down quark. Note that anti-commuting the left/right handed projector through the γ_0 converts P_+ into P_- and P_- into P_+; also observe that $P_+P_- = 0$ while $P_+P_+ = P_+$ and $P_-P_- = P_-$. The chiral Fermion Lagrangian is then,

$$\begin{aligned}
\hat{\mathcal{L}}_{\text{Fermion}} = \hat{\bar{E}}_L \not{D}_{L,l}\hat{E}_L + \hat{\bar{E}}_R \not{D}_{R,l}\hat{E}_R + \hat{\bar{Q}}_L(\not{D}_{L,q} + \alpha_S\gamma^\mu C_\mu)\hat{Q}_L \\
+ \hat{\bar{Q}}_R(\not{D}_{R,q} + \alpha_S\gamma^\mu C_\mu)\hat{Q}_R. \quad (675.1)
\end{aligned}$$

I have added in the strong interaction gauge field C_μ for the quark fields.

Can we also have a Boson field? Yes, we can, and I will call it H. H has dimensions $[x]^{-1}$, which means that the simplest allowed action for H is

$$\hat{\mathcal{L}}_H = H^\dagger(i\partial_\mu + \alpha_W\tau^a B^a_\mu + \alpha A_\mu)(i\partial_\mu + \alpha_W\tau^a B^a_\mu + \alpha A_\mu)H + \frac{\lambda}{4}(H^\dagger H)^2 - \frac{1}{2}\mu^2(H^\dagger H). \quad (675.2)$$

λ is another dimensionless parameter, and μ is a dimensional parameter. This is the most general expression possible for a scalar (no direction index) Boson field consistent with the symmetries and requirement for gauge invariance. I have written B_μ as a sum over the three individual gauge fields proportional to matrices τ^1, τ^2 and τ^3 which are the three anti-commuting Hermitian and traceless 2×2 matrices. I will take μ to be the one dimensional parameter I am permitted in the theory without making the theory scale dependent. I have chosen to couple the H field to the A_μ and B_μ gauge fields but not the C_μ gauge field, and it can also couple with the Fermion fields. This means that H must in practice be either a two component complex vector or a 2×2 Hermitian matrix (since B_μ, a 2×2 matrix, acts upon it). Both of these are described by 4 variables, the same number that we have in the gauge fields. Conventionally, one treats H as a 2-component vector (as I did in chapter 10, but

here I will use the matrix formulation since it is one way of extending the standard model to satisfy the experimental result that the neutrinos have masses. I will write the Boson field as a unitary matrix U times a Hermitian Matrix $\hat{H}$. We are going to want to couple this to the left and right handed Fermion fields using an expression like $\hat{\bar{\psi}}_L U \hat{H} \hat{\psi}_R$, although for the moment I will just consider the theory with this H field and the A_μ and B_μ gauge fields. The unitary matrix is affected by the SU(2) gauge transformation on the left-handed fields and the U(1) transformation; the $\hat{H}$ matrix varies under the U(1) gauge transformation of the right handed fields. Being gauge dependent, the matrix U is not observable. This means that the gauge covariant derivative of the Bosonic field gives us

$$D_\mu[U\hat{H}] = U(U^\dagger \partial_\mu(U) + \partial_\mu \hat{H}) = U(\partial_\mu \hat{H} - i(\alpha_W \tau^a B^a_\mu + \tau^0 \alpha A_\mu)\hat{H}), \tag{676.1}$$

and the kinetic term in the action has to be proportional to,

$$\begin{aligned} &\mathrm{tr}((D_\mu[U\hat{H}])^\dagger D_\mu[U\hat{H}]) = \\ &\qquad \mathrm{tr}(\partial_\mu \hat{H} - i(\alpha_W \tau^a B^a_\mu + \tau^0 \alpha A_\mu)\hat{H})(\partial_\mu \hat{H} - i(\alpha_W \tau^a B^a_\mu + \tau^0 \alpha A_\mu)\hat{H}), \end{aligned} \tag{676.2}$$

with,

$$\begin{aligned} \tau^1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \tau^2 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \\ \tau^3 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \tau^0 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{676.3}$$

Meanwhile, the Boson field also has a potential term, which we can write as,

$$-\frac{\mu^2}{2}\mathrm{tr}(\hat{H}^2) + \frac{\lambda}{4}(\mathrm{tr}(\hat{H}^2))^2,$$

which with $H = \hat{H}_a \tau^a$ has $\mathrm{tr}(\hat{H}^2) = \hat{H}_0^2 + \hat{H}_1^2 + \hat{H}_2^2 + \hat{H}_3^2$. The total Lagrangian is

$$\begin{aligned} \hat{\mathcal{L}}_{\mathrm{Higgs}} = &\mathrm{tr}(\partial_\mu \hat{H} - i(\alpha_W \tau^a B^a_\mu + \tau^0 \alpha A_\mu)\hat{H})(\partial_\mu H - i(\alpha_W \tau^a B^a_\mu + \tau^0 \alpha A_\mu)\hat{H}) - \\ &\frac{\mu^2}{2}\mathrm{tr}(\hat{H}^2) + \frac{\lambda}{4}(\mathrm{tr}(\hat{H}^2))^2 + F^2_{\mu\nu}[A_\mu] + F^2_{\mu\nu}[B_\mu] \end{aligned} \tag{676.4}$$

The trace is over the components of $\hat{H}$. For the more complex B_μ matrix, the gauge field action takes a slightly different form,

$$F_{\mu\nu}[B_\mu] = \partial_\mu B_\nu - \partial_\nu B_\mu + \frac{i}{\alpha_W}[B_\mu, B_\nu]. \tag{676.5}$$

That final term means that there can be couplings between 3 or 4 B gauge fields, i.e. we can have a situation where a B_μ field decays into two B_ν fields, or a B_μ field decays into a B_μ field and two B_ν fields. This self-interaction makes the more complex matrix gauge interactions (called non-Abelian gauge theories) more complex than electromagnetism (called an Abelian gauge theory), and they have very different properties. The gauge transformation rule for B_μ in a non-Abelian theory is also slightly different to the simple Abelian case.

$$B_\mu \to B_\mu + \frac{1}{\alpha_W}\partial_\mu \alpha + i[\alpha, B_\mu], \tag{676.6}$$

where α is the 2×2 matrix that parametrises the unitary rotation of the Fermion or H fields.

Going back to the Lagrangian $\hat{\mathcal{L}}_{\mathrm{Higgs}}$, we see that the minimum action for the H field is not at a constant $\hat{H} = 0$, but instead $\hat{H} = H_M + \hat{H}'_a \tau^a$ $(a = 0, 1, 2, 3)$, where $\mathrm{tr}H^2_M = \mu^2/\lambda$. One possible solution of this is

$$H_M = \frac{\mu}{\sqrt{\lambda}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{676.7}$$

although there are, of course, many other possibilities. That the minimum of the action is not at $\hat{H} = 0$ is significant, because the minimum action solution corresponds to the classical solution, which is the most important contribution to the path integral. When we perform an expansion, we take the most important contribution as the basis of the expansion, and then compute corrections to it; simply because we want those corrections to decrease as we improve the approximation so we start with the most important term. So this means that we instead expand around this minimum action solution. Before we start, we have to choose one ground state to be the basis of the expansion. I will select the state H_0 above, but that we have to make this choice is significant, because by doing so we break the $U(2)$ symmetry. This symmetry breaking is not at the level of the action, but it appears in the solution, a phenomenon known as spontaneous or hidden symmetry breaking, and which plays a crucial role in modern physics. So we write $\hat{H}$ as $H_M + \hat{H}'$, and immediately we see that something strange happens to the Lagrangian:

$$\hat{\mathcal{L}}_{\text{Higgs}} = \text{tr}\Big[(i\partial_\mu + \alpha_W \tau^a B^a_\mu + \alpha A_\mu)(H_M + \hat{H}'_a\tau^a)(i\partial_\mu + \alpha_W \tau^a B^a_\mu + \alpha A_\mu)(H_M + \hat{H}'_a\tau^a) + F^2_{\mu\nu}[A_\mu] + F^2_{\mu\nu}[B_\mu] - \frac{\mu^4}{4\lambda} + \mu^2(\hat{H}'_-)^2 + \frac{\lambda}{4}((\hat{H}'_1)^2 + (\hat{H}'_2)^2 + (\hat{H}'_+)^2 + (\hat{H}'_-)^2)^2 + \mu\sqrt{\lambda}\hat{H}'_-((\hat{H}'_1)^2 + (\hat{H}'_2)^2 + (\hat{H}'_+)^2 + (\hat{H}'_-)^2)\Big], \quad (677.1)$$

with $\hat{H}'_+ = \frac{1}{\sqrt{2}}(\hat{H}'_0 + \hat{H}'_3)$ and $H'_- = \frac{1}{\sqrt{2}}(\hat{H}_0 - \hat{H}_3)$.

Every quadratic term in the fields with no derivatives represents a mass. What we see is that the $\hat{H}'_-$ field has a mass μ^2 and the remaining three fields have no mass. This would normally lead to an extra three massless particles (which we don't observe in nature), but in practice in this particular case they can be removed by performing a gauge transformation: they are absorbed into the gauge fields. We are left with one massive particle, whose creation and annihilation operators are contained within $\hat{H}'_-$, known as the Higgs Boson.

But equally interesting is that the gauge field also acquires a mass term: we have within the Lagrangian the term

$$H_M(\alpha_W \tau^a B^a_\mu + \alpha\tau_0 A_\mu)^2 H_M. \quad (677.2)$$

Normally, gauge symmetry prevents a mass term arising for the gauge fields at the level of the Lagrangian, and that is still the case. However, this electroweak spontaneous symmetry breaking allows the gauge fields to appear as though they are massive particles.

In practice, there are three massive gauge fields and one massless field. The massive fields and their masses m_W and m_Z are

$$\begin{aligned} W^+_\mu &= \frac{1}{\sqrt{2}}(B^1_\mu + iB^2_\mu) & m_W &= \alpha_W \frac{\mu}{2\sqrt{\lambda}} \\ W^-_\mu &= \frac{1}{\sqrt{2}}(B^1_\mu - iB^2_\mu) & m_W &= \alpha_W \frac{\mu}{2\sqrt{\lambda}} \\ Z^0_\mu &= \cos\theta_W B^3_\mu + \sin\theta_W A_\mu & m_Z &= \sqrt{\alpha_W^2 + \alpha^2}\frac{\mu}{2\sqrt{\lambda}}, \end{aligned} \quad (677.3)$$

while the massless field is

$$\Gamma_\mu = -\sin\theta_W B^3_\mu + \cos\theta_W A_\mu, \quad (677.4)$$

and $\sin\theta_W = \frac{\alpha}{\sqrt{\alpha_W^2+\alpha^2}}$. The $W^\pm$ fields are re-arranged from the B^1 and B^2 fields so they carry precise electric charge. The W^+ field, for example, couples the anti-electron with the neutrino, so an anti-electron can decay into a W^+ and neutrino, or an electron into a W^- and an anti-neutrino. They couple with Γ in the new expression of the Lagrangian, meaning

that they can emit or absorb the massless particles (which are known as photons), while the Z^0 field does not emit or absorb photons.

Now its time to go back to the Fermion fields. Earlier we used the Dirac operator,

$$\not{D} = \gamma^\mu(i\partial_\mu + Q\alpha A_\mu + Q_W\alpha_W B^a_\mu \tau^a), \tag{678.1}$$

which in terms of the new fields becomes

$$\not{D} = \gamma^\mu(i\partial_\mu + \alpha_W Q_W(W^+_\mu\tau^+ + W^-_\mu\tau^-) + \frac{\alpha_W}{\cos\theta_W} Z^0_\mu(Q_W\alpha_W\cos\theta_W\tau^3 - \alpha\sin\theta_W Q) + (Q_W\tau^3 + Q)e\Gamma_\mu \tag{678.2}$$

with $\tau^\pm = \frac{1}{\sqrt{2}}\tau^1 \pm i\tau^2$ and $e = \alpha\alpha_W/\sqrt{\alpha^2+\alpha_W^2}$. In nature, the left handed fields are coupled to the full U(2) symmetry group (i.e. both B and A), while the right handed fields just transform according to a $U(1)\times U(1)$ symmetry group (i.e. just the B^3 and A fields), which can be seen as two independent electromagnetic rotations. For the left handed Fermions we have $Q_W = -\frac{1}{2}$ and $Q = -\frac{1}{2}$ for the leptons and $Q = \frac{1}{6}$ for the quarks. Since Γ_μ is the photon field which conveys the electromagnetic force, this gives an electric charge of -1 for the electron, 0 for the neutrino, $\frac{2}{3}$ for the up quark, and $-\frac{1}{3}$ for the down quark. For the right handed fields, we have $Q_W = 0$ and we are now free to choose Q independently for each type of particle, giving the same electric charges. Note that as charged particles, W^+ and W^- are also subject to the electromagnetic gauge transformations.

Because we have turned the left and right handed Fermions into different types of particles, and in particular subject to different gauge symmetries, we can't introduce a mass in the natural way of $\hat{\bar{E}}_{La}M_{ab}\hat{E}_{Rb}$. But we do have the possibility of couplings between the Fermion fields and the Boson field,

$$\hat{\mathcal{L}}_{fH} = \lambda_e\hat{\bar{E}}_R H\hat{E}_L + \lambda_e\hat{\bar{E}}_L H\hat{E}_R + \epsilon^{ab}\epsilon^{cd}\lambda_{\nu_e}(\hat{\bar{E}}_{Ra}H_{bd}\hat{E}_{Ld} + \hat{\bar{E}}_{La}H_{bd}\hat{E}_{Rd}) + \lambda_d\hat{\bar{Q}}_R H\hat{Q}_L + \lambda_d\hat{\bar{Q}}_L H\hat{Q}_R + \epsilon^{ab}\epsilon^{cd}\lambda_u(\hat{\bar{Q}}_{Ra}H_{bd}\hat{Q}_{Ld} + \hat{\bar{Q}}_{La}H_{bd}\hat{Q}_{Rd}), \tag{678.3}$$

where ϵ is the anti-symmetric tensor ($\epsilon^{12} = 1, \epsilon^{21} = -1, \epsilon^{11} = \epsilon^{22} = 0$), a, b, c and d can take the values 1 or 2, and $\hat{\bar{E}}_{R1}$ means the first component of $\hat{\bar{E}}_R$ while $\hat{\bar{E}}_{R2}$ means the second component. λ_e, λ_{ν_e}, λ_u and λ_d are dimensionless constants, which will ultimately give the masses of the various particles. So in the unbroken Lagrangian, the Fermions are all massless. However, after electroweak spontaneous chiral symmetry breaking, they each acquire a mass, for example $\lambda_e\frac{\mu}{\sqrt{\lambda}}$ for the electron.

The choices of electric charges above might seem strange, but actually they are required for the self consistency of the theory. There is an issue that only the left handed fields see the SU(2) gauge transformation, but the right handed fields don't. Usually, when the measure of the path integral – the $\int d[\psi]d[\hat{\bar{\psi}}]$ term that defines how we add up all the different paths transforms under a gauge transformation, the contributions from the left and right handed Fermion fields cancel out, leaving the measure gauge invariant[6]. But in this case, only the left handed fields transform under the gauge transformation. The measure of the path integral is not unchanged under the gauge symmetry, and since the measure is part of the description of how God sustains the universe, it should respect the symmetry that because God is outside space and time there is no preferred gauge. The measure should, if it is to be consistent with our premises, be unchanged under a gauge rotation, but once we treat the left and right handed Fermion fields differently it isn't; this is known as the anomaly. It is a serious problem, unless, of course, there is a miraculous cancellation between the contributions from all the different Fermion fields. The charges we noted before were,

[6]This is not quite correct – there is a contribution called the topological charge, but I will neglect this here.

	e_L	ν_L	u_L	d_L	e_R	ν_R	u_R	d_R
Q	-1/2	-1/2	1/6	1/6	-1	0	2/3	-1/3

and it happens that these are precisely the combination needed to cancel out the various anomalies. Without this cancellation, the theory would be inconsistent. Thus this precise choice of electric charges is needed to maintain the self-consistency of the theory (given that the weak interaction only affects left handed Fermions).

We are almost finished. So far, we have just described one family of Fermion particles. There is no reason why there should just be one family of particles, and in practice there seems to be three. This leads to an additional complication. It allows us to perform another rotation of the Fermion fields, this time with respect to the flavour indices. Again, we need not do this, but there is no reason why we should not , and in practice nature does. We add an index f to the $\hat{\bar{Q}}$ and $\hat{Q}$ operators, and rotate them using the matrix

$$\begin{aligned}\hat{Q}_{Lf} &\to P_+ \begin{pmatrix} U_{uff'} & 0 \\ 0 & U_{dff'} \end{pmatrix} \begin{pmatrix} \hat{u}_{f'} \\ \hat{d}_{f'} \end{pmatrix} \\ \hat{Q}_{Rf} &\to P_- \begin{pmatrix} U_{uff'} & 0 \\ 0 & U_{dff'} \end{pmatrix} \begin{pmatrix} \hat{u}_{f'} \\ \hat{d}_{f'} \end{pmatrix}. \end{aligned} \tag{679.1}$$

U_u and U_d are unitary rotations which act on the f or flavour indices. They act between different types of quarks. This rotation only affects the part of the Lagrangian involved in the Weak interaction (the W^+ and W^- gauge Bosons interact with an up quark and down quark and thus change quark flavour; the other gauge Bosons don't change flavour), where it introduces a matrix $V = U_u^\dagger U_d$ into the Lagrangian, which allows mixing between the different quark flavours and different families of quarks. This matrix is known as the Cabibbo-Kobayashi-Maskawa matrix, and once we have removed the various redundant variables it is described by four parameters. A similar matrix can exist in the electron/neutrino sector and allow oscillations between the various neutrinos. Note that the right handed neutrino[7] has no electric charge, and no coupling to the weak or strong nuclear forces. It can only interact with other matter through the Higgs particle and gravity.

If we put all these together, we find that the Lagrangian of nature is (equations (674.4), (676.3), and (673.1))

$$\hat{\mathcal{L}} = \hat{\mathcal{L}}_{\text{Higgs}} + F_{\mu\nu}[C_\mu]^2 + \hat{\mathcal{L}}_{fH} + \hat{\mathcal{L}}_{\text{Fermion}}. \tag{679.2}$$

And that is it. The terms quadratic in the fields describe propagators, the probability amplitude for a particle to travel from one place to another. The terms with more than three fields in the Lagrangian describe vertices, giving the probability amplitudes for various particles to be created and destroyed, and consequently the laws of efficient and final causality for each type of fundamental particle. That's the standard model. The Lagrangian is a little more complicated than we would like (and, of course, we still have to renormalise it), but in four equations we have described everything in the universe except gravity.

And I do mean everything. Let us suppose that we can write down an equation which combines (677.2) with a quantum theory of gravity. Everything in the material universe would be described by the resulting equation. All of biology, all of chemistry, all of the

[7]Note again that the existence of the right handed neutrino is only speculative. It is not part of the conventional standard model of particle physics. Some extension to the standard model is required to account for the neutrino masses, and what I have proposed here is just one of a number of possibilities for this extension, and as yet which of these possibilities is correct is still unknown. The nature of the electromagnetic, strong, and weak nuclear forces are firmly established by experiment, at least in the limit of small gravitational field. The existence of the Higgs Boson is confirmed, but its nature is still being investigated as I write this. The particle found seems to be the standard model Higgs, consistent with the matrix formulation I have outlined above and the more conventional complex 2-vector formulation. Those parts of my description are correct. It is not yet known whether there are additional particles on top of these, or quite how the neutrino acquires its mass. The existence of right handed neutrinos, which my expression of the model requires, is unproven and could be false.

physics of astronomy, motion, optics, fluids, solids, the weather, nuclear and atomic physics would all be described by that one equation. (677.2) is almost there. It looks a bit ugly (once we substitute in the definitions for $\mathcal{L}_{fH}$ and so on), but considering the vast range of phenomena it describes, it is remarkably concise.

Of course, extracting all that science from this one equation is by no means easy, but we have made good progress, especially with the electromagnetic and weak interactions where we have an established method to solve the equations. The strong interaction is a bit harder, because the standard perturbative series doesn't converge and the vacuum state isn't the $|0\rangle$ state we use in electromagnetism but contains complicated structures made from the gluon (C_μ) fields, and the vacuum of the strong interaction is still not well understood[8]. However, even here, we are making a great deal of progress through various numerical calculations, models, and the perturbative calculations we can do. ■

Discussion:

I should note that my expression for the standard model Lagrangian here is a little non-standard; usually the H field is set up as a two component vector rather than a 2×2 matrix, which means that equation (676.3) is expressed a little differently. The right handed neutrino (which doesn't interact with any of the gauge Bosons) is excluded from the standard model, and the neutrinos are in the standard formulation massless. Experiments have now made it certain that the neutrinos do have a small mass, which means that some correction to the standard model is needed. I have chosen what seems to me to be the simplest modification to the standard model which gives us neutrino masses; an amendment to this work will probably be required once we know what the real answer is. This amendment, however, will not affect my overall approach: all proposed models of neutrino masses involve quantum field theories subject to the symmetries which are required by my axioms.

I should also mention that I have over-simplified this discussion in several places; there are numerous details which I have either omitted or expressed too concisely. My calculation is not as rigorous as I would make it were I writing a physics textbook: my purpose has been to give a general idea of how we can achieve a profound result from just a few premises; to do the calculation fully would require even denser mathematics than I have allowed myself here. Nor have I described how one conducts calculations in the standard model, both in perturbation theory and otherwise, nor the important topological properties of field theory, nor topics such as asymptotic freedom and how the renormalised coupling constant changes with the energy scale, nor conservation laws and conserved currents, nor the topological aspects of the gauge field 'vacuum', nor anomalies and symmetry breaking by the measure of the theory, nor spontaneous symmetry breaking or the accidental discrete symmetries of the theory, nor how this theory can be extended into nuclear physics, atomic physics and subsequently chemistry. In any complete discussion of the physics these topics and more would have to be addressed. However all these topics follow from the symmetry considerations and the Lagrangian (or are straight-forward extensions of it), so they do not modify the overall philosophy.

It is also worth listing the particles of the standard model. For the Fermions we have three electron type quarks (known as the electron, muon, and tauon); three types of neutrino (electron neutrino, muon neutrino, and tau neutrino), six quarks (up, down, strange, charm, top, bottom) which come in three different colours each for a total of eighteen quarks. There are also the corresponding anti-particles, and each of these particles and anti-particles have two spin states. For the gauge Bosons, we have the photon which carries the electromagnetic force, the three weak nuclear force carriers, the $W^\pm$ and Z^0, and the eight gluons described by the C_μ gauge field. Finally, there is the remaining Higgs particle generated by $\hat{H}_-$. I apologise for the names given to these particles. Up to the early twentieth century, physicists

[8]Newton could not solve the three body problem; modern physicists struggle with the zero body problem. That's progress for you.

were well versed in the classical languages, so invented nice classical names such as 'electron' for the particles they discovered. But from the mid-twentieth century, physics became a lot less sophisticated, and so we are stuck with names like 'quark' (a type of cream cheese popular in Germany; though historically taken from a nonsense poem by James Joyce) 'strange' and (worst of all) 'glue-on'. If I could change these names, I would, but like the rest of the physicist community, I'm stuck with them.

We should also look at the various points where God could have constructed the universe differently. Firstly, we have the choice of three space and one time dimensions. There is no obvious reason from the underlying philosophy why the universe has to be like this. Then there is the choice of a $U(1) \otimes SU(2) \otimes SU(3)$ gauge symmetry - a $U(1)$ gauge symmetry I think is necessary if we are to avoid a wholly static universe, but why the other gauge groups, and why only these gauge groups? Why do leptons interact with the $U(1) \otimes SU(2)$ gauge Bosons, and quarks with all three? Why three families of particles[9]? Why does the weak interaction only interact with left handed particles? Why only quarks and leptons? Why does the Higgs couple to the $U(1) \otimes SU(2)$ sector of the gauge fields, but not the $SU(3)$ sector? Why do the possible $F^{\mu\nu}F^{\rho\sigma}\epsilon_{\mu\nu\rho\sigma}$ terms not appear in the Lagrangian? These are all choices which God made, but did not have to; however, without them the universe would be a very different place. But then we have the most mysterious thing of all: the various parameters of the theory. I'll list them out. We have α, giving the strength of the electromagnetic force, α_S giving the strength of the strong interaction, and the angle θ_W relating the strength of the electromagnetic and weak forces; then we have λ, the self interaction of the full matrix of the H Boson, the twelve couplings between the Fermions and H which lead to the particle masses. The Cabibbo-Kobayashi-Maskawa matrix has an additional four parameters, and there is the θ_2 coupling in the strong nuclear theory for the $F_{\mu\nu}F_{\rho\sigma}\epsilon^{\mu\nu\rho\sigma}$ term. That makes a total of twenty one numbers which describe the standard model of particle physics. There is no theoretical reason (at least that we know of so far) why they should take one value rather than another. We can only measure them through experiment. Finally, there is μ, related to the mass of the Higgs particle, which, as a dimensional quantity, depends on our choice of units, and provides a natural physical scale. Other dimensional quantities include the speed of light c and Planck's constant $\hbar$; as with μ these should, however, be thought of as definitions of the units we use to measure time and energy and as such should not be thought of as physical constants.

Is there any more to particle physics than this? We haven't yet found anything. There might be something else out there, or there might not. Almost certainly, there will be. Something is needed to merge gravity into the standard model, and here currently we have a lot of theories but no hard facts. The issue of dark matter and dark energy from cosmology need explaining. The neutrino masses are also a mystery, although hopefully one that will be resolved soon. □

Now, of course, very little of what I have written here is new. Particle Physics theorists would have recognised it as a (very much non-standard) derivation of Richard Feynman's path integral formulation of the standard model of particle physics (Feynman and Hibbs,

[9]This is actually significant – in the standard model, there is a slight violation of the symmetry between particles and anti-particles due to a term in the Cabibbo-Kobayashi-Maskawa matrix, which helps explain why the universe is matter rather than anti-matter dominated (though the conventional wisdom is that it is not sufficient to explain the matter dominance we observe in nature). This violation is only possible if there are at least three families of Fermions. A fourth family is not completely ruled out, but seems very unlikely. For example, the experimental decay width of the Z^0 Boson can be accounted for if we include three families of neutrinos; a fourth neutrino would lead to a different theoretical prediction which would disagree with experiment, unless the mass of that fourth neutrino was of the same order of the Z^0 and a decay into that particle were forbidden by the conservation of energy. This seems unlikely, since the neutrinos are by far the lightest particles with masses very close to zero, and we can expect the same for a hypothetical fourth neutrino. Of course, 'expect' is not the same as a firm proof, but in light of this and other experimental evidence, a fourth family of Fermions would be an enormous shock to put it mildly.

1965). I have simplified on a few points, skipped over a lot of details, some of my notation is a little non-standard, and I have taken a slightly different path to the standard one in the early stages of my argument, but the bulk of this argument is just repeating material which can be found in any advanced particle physics textbook (only there the argument would be more detailed and with considerably more mathematical rigour). The only novelty of any substance I have added to the argument is my explanation of the origins of the symmetry laws and the path integral formulation itself in terms of classical theism. In the physics textbooks, these are just taken for granted. For example, by the time physicists have reached this stage of their education, we have all learnt about gauge invariance from classical electrodynamics, and how it arises from experiment, and it just seems natural to put it into the theory here. Similarly the path integral formulation originally arose from Feynman's thoughts about the standard two slit experiment in quantum mechanics (or classical optics) – in other words, it was ultimately inspired by experiment. I do not believe that the physicist working as a physicist can ultimately answer the question of why the universe has these symmetries, or why this path integral formulation works so well. This is the task of the philosopher (or the physicist pretending to be a philosopher). Granted the symmetry requirements seem obvious and natural to me, whether we are discussing some mechanical laws of physics (immutable, unchangeable in space and time) or a theological system such as I have developed here. But why the path integral? Why can the universe jump from one state to any other (baring the symmetry requirements) at any time? This is certainly inconsistent with the naive mechanism which motivated the likes of Richard of Wallingford, Descartes, Galileo, Newton and Kant. That the universe seems to be directed by a free will ought to be troubling to any atheist.

I have shown that the basic form of the standard model can be extracted from the basic principles of classical theism – although I had to make a number of seemingly arbitrary choices along the way to get to our particular universe. It may, of course, be that it can also be predicted from various atheistic metaphysics – I cannot disprove that, though it would be a very different form of atheism to those promoted by modern philosophers. But for the purposes of this work, I don't need to. My goal is to show that modern physics is consistent with classical theism; and if my argument so far stands up, then that is what I have done, at least for the standard model of particle physics.

15.9. Gravity

So what of the other leg of modern physics, general relativity? Classical general relativity is, of course, seen as the final triumph of the mechanistic philosophy. If, however, I can show that gravity is not in fact described by classical general relativity, but by some quantum theory – probabilistic rather than deterministic – then I would have completed my case. Nothing would remain of the physics that was inspired by the mechanical philosophy, and there would then be no reason to maintain the mechanical philosophy as part of our underlying metaphysics. And those forms of deism and atheism which depend on it would collapse with it.

Now, it is widely believed that general relativity should be replaced by some quantum theory of gravity, but now we come to the biggest challenge: it is not yet known what the quantum theory of gravity is. Thus I am now forced to move from what is experimentally secure to what is highly uncertain. Nonetheless, I will proceed as well as I can, with the proviso that my conclusions might need to be modified if physics develops in a different direction to how I suspect it will. I should also say that the approach I discuss here is even more speculative than most discussions on this topic: rather than describing one of the mainstream approaches to quantum gravity such as string theory or loop quantum gravity, what I want to present is my own tentative ideas and research on the topic. I still believe that the more standard approaches to quantum gravity support my overall thesis, namely that the true philosophy of physics is closer to the classical Aristotelian picture than any

of the ideas that were spawned by the enlightenment. However, the approach I adopt here puts gravity on the same footing as much of the rest of modern physics. It makes it simply a matter of respecting a symmetry. Thus the link between gravity and metaphysics remains clear and obvious. I should also apologise that again my discussion will be rather technical.

Firstly, I need to discuss why we believe that classical gravity is inadequate; so I need to quickly review the classical theory of gravity.

There are five key components in general relativity (see section 7.7).

(1) Firstly, we have the metric tensor $g_{\mu\nu}$.

(2) Secondly, the coordinate transformation $\frac{\partial x'^{\mu}}{\partial x^{\nu}}$.

The metric tensor is used to create a scalar (and thus coordinate system invariant quantity) out of two four vectors. So, for example, the path of light which originated at a point x_0 is described in any coordinate frame by the equation $\sum_{\mu\nu}(x^{\mu}-x_0^{\mu})g_{\mu\nu}(x^{\nu}-x_0^{\nu})=0$, where x corresponds to the path of the light and x_0 is some fixed point of reference. In flat space time (and, indeed, the standard model of particle physics), the metric is diagonal, with elements $(-1,1,1,1)$ (using units where the speed of light is one). When we switch to a different coordinate system, say x', we use the transformation matrix, $(x'^{\mu}-x_0'^{\mu})=\frac{\partial x'^{\mu}}{\partial x^{\nu}}(x^{\nu}-x_0^{\nu})$ (technically this expression applies when x is very close to X_0). At the same time, we modify the metric according to $g'_{\alpha\beta}=g_{\mu\nu}\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}}$. Thus the scalar product $(x^{\mu}-x_0^{\mu})g_{\mu\nu}(x^{\nu}-x_0^{\nu})$ remains the same no matter which coordinate system we use.

In general relativity, the primary focus is on local rather than global coordinate transformations. A global transformation is where the same thing happens everywhere in the universe. This is what is covered by special relativity, and is already included in the standard model. A local transformation is when we focus on each point individually, and transform them in different ways. We focus on those places were we are in the immediate vicinity of the reference point, and then build up the global picture one step at a time rather than all at once.

(3) Next, we need the derivative operator ∇_{μ}, describing the variation of whatever is being differentiated under infinitesimal changes of the new coordinates x'. We choose to consider that derivative operator where the differential of the metric is zero: $\nabla^{\mu}g_{\mu\nu}=0$. This derivative operator is linked to the derivative in the original x coordinates by the Christoffel symbol Γ, so $\nabla_{\mu}t^{\nu}=\partial_{\mu}t^{\nu}+\Gamma^{\nu}_{\mu\rho}t^{\rho}$ for some four vector t. The Christoffel symbol can be expressed in terms of the metric tensor and tits derivatives with respect to the original coordinate x. We also need the concept of curvature. The curvature tensor $R_{\mu\nu\rho}{}^{\sigma}$ can also be expressed as the commutator between two derivative operators acting on a vector: $R_{\mu\nu\rho}{}^{\sigma}t_{\sigma}=(\nabla_{\mu}\nabla_{\nu}-\nabla_{\nu}\nabla_{\mu})t_{\rho}$. In flat space time, this curvature tensor is zero, but in curved space time it is non-zero (hence the name of the tensor). It is basically a measure of how far away we are from a Euclidean or hyperbolic geometry.

(4) The fourth component of general relativity is the stress energy tensor, $T_{\mu\nu}$. This is a 4×4 real symmetric matrix, so has ten independent components. Its time-time component gives the energy density (i.e. the energy contained within a unit of volume), its three time-space components give a momentum density, and its six space-space components describe the sheer stress and pressure of whatever matter is described by the tensor.

Now this object can be constructed in two distinct ways, which is going to be of importance to us later. The first way is as the Noether current associated with gauge covariant time or space translations. Using gauge covariant translations (i.e. to perform a gauge transformation at the same time as we translate) is important – you will not get the right answer without it. You will recall that a spinor field can be freely rotated by a unitary matrix $\psi(x)\to U\psi(x)$. To

ensure that the derivative of the spinor retained this property, we modified it so that $\frac{\partial}{\partial x^\mu} \to \frac{\partial}{\partial x^\mu} - igA_\mu$, with $igA_\mu = U^\dagger \partial_\mu U$. We need the same trick here. Normally, a small translation $x^\mu \to x^\mu + \epsilon^\mu$, with ϵ a sufficiently small number that we can neglect any terms of order ϵ^2, can be generated using the operator $1 + \epsilon^\mu \frac{\partial}{\partial x^\mu}$. So, for example, $\psi(x + \epsilon) = (1 + \epsilon^\mu \frac{\partial}{\partial x^\mu})\psi(x) + O(\epsilon^2)$. To make a gauge covariant derivative operator, we just replace this derivative with the derivative operator which includes the gauge field. The generator of the translation is $\psi(x+\epsilon) = (1+\epsilon^\mu(\frac{\partial}{\partial x^\mu} - igA_\mu))\psi(x) + O(\epsilon^2)$. The corresponding transformation of the gauge field is $A_\mu \to A_\mu + ig\epsilon^\nu F_{\nu\mu}$. It can be seen that the standard model action remains the same if this transformation is applied and ϵ^μ does not vary across space – ϵ^μ is a constant, independent of the position x^μ. If we consider local coordinate transformations, i.e. ϵ^μ is a function of space, then we find that the variation of the action under the transformation is proportional to $\partial^\mu \epsilon^\nu T_{\mu\nu}$, where $T_{\mu\nu}$ is the stress energy tensor. Now if we demand that no physical observable depends on which coordinate system we use, then we require that this change to the action is zero, which in effect requires that $\partial^\mu T_{\mu\nu} = 0$. This is, of course, satisfied in flat space time. In curved space time, the right hand side of this equation is proportional to the derivative of Einstein's curvature tensor. Remember, the partial derivative ∂^μ is a measure of change, and thus $\partial^\mu T_{\mu\nu} = 0$ is a statement that the total amount of energy and momentum in the universe doesn't change. In classical physics, this is the basis of the conservation of energy and momentum.

The second way to derive the stress energy tensor, is simpler, and far more commonly used in the literature on general relativity. The stress energy tensor is the derivative of the action with respect to the metric. This gives something proportional to the result obtained from the symmetry. Why is it preferred? One reason is that the standard model action is also invariant under the global 'translation' defined with just the normal derivative operator. In practice, this transformation is actually the combination of a global translation and a global gauge transformation, but since the action is invariant under each of these, and the global transformations commute (i.e. it doesn't matter which order you apply them in), then the action would also obviously be invariant under their combination. However, when we consider a local coordinate transformation and a local gauge transformation, it does matter which order you apply these in, so you can't apply either transformations independently. If people try to do so, then they will get an obviously wrong Noether current (the stress-energy tensor is the Noether current of the symmetry). However, the lure of the global transformation has proved too much for many people, and they have naively converted it into a local transformation, and called the nonsense result as though it were the Noether version of the stress energy tensor. This differs from the gravitational version (which isn't nonsense) which is derived from the derivative of the action with respect to the metric. With one method seemingly giving the wrong answer, and the other method giving the right answer, it is natural to prefer the method which gives the right answer. Of course, the first method gave the wrong answer because it was utilised incorrectly. But many people did not notice that.

Secondly, the Einstein-Hilbert action was found, which led to Einstein's field equation upon minimisation with respect to variations in the metric. It could generate the equations for general relativity in the same way that a classical action can generate Newton's equations of motion and Newton's law of gravity. But only if we treat the metric as a dynamical field in the same way that the gauge potential is treated in the action for electro-magnetism. This led to the natural idea that the metric itself plays the same role as the fields of classical (or even quantum) electrodynamics. This idea moves us away from the idea of understanding general

relativity as a consequence of a symmetry principle, and instead places it in the same class of field theories as govern the rest of physics. The two approaches are related two different approaches to the philosophy of general relativity. The first sees it from a more mathematical geometrical perspective; the second a more physics based mechanical perspective (where interactions are either the consequences of forces or the exchange of particles).

(5) Finally, we need the notion of a geodesic. This is the equivalent of a straight line in standard flat coordinates; it represents a curve whose tangent vectors remain parallel as they are transported along it.

General relativity is based on two principles: firstly, that objects travel along geodesics unless acted on by an external force, and secondly Einstein's field equation, which connects the curvature (and thus the metric) with the stress-energy tensor:

$$R_{\mu\rho\nu\sigma}g^{\rho\sigma} - \frac{1}{2}g^{\alpha\beta}g^{\rho\sigma}g_{\mu\nu}R_{\alpha\rho\beta\sigma} + \Lambda g_{\mu\nu} = -8\pi G T_{\mu\nu}. \tag{685.1}$$

Λ is known as the cosmological constant, and G is Newton's constant. Einstein's field equation can be obtained from a least action principle, by adding the term $\sqrt{\det g}g^{\alpha\beta}g^{\rho\sigma}R_{\alpha\rho\beta\sigma}$ and minimising the total action with respect to variations in the metric g. This equation is short, easy to write down, but contains a lot of physics. Hidden within these two equations lie buried an updated version of Newton's laws of motion, gravity, the evolution of galaxies and the solar system, the expansion of the universe, the big bang, black holes, and everything related to cosmology baring the first short period of the universe and the finer details of black hole physics.

The standard model of cosmology, which adds cosmic inflation[10], the big bang[11], and dark matter and energy[12] to one particular solution of Einstein's field equation, ranks alongside the standard model of particle physics as one of the best experimentally attested and most powerful scientific theories we know of. Aside from the big bang, which follows inexorably from Einstein's theory of gravity, what these additional ideas we need to add to general relativity are is not clear; that something is there, and we know a few general properties of that something, is pretty much certain from experimental observations combined with our best theory.

Now we are almost in a position to see the difficulty in combining general relativity with quantum field theory. Our goal is to find a quantum theory of gravity, which a) is self consistent; b) reduces to Einstein's General relativity in the classical $\hbar \to 0$ limit; and c) reduces to the standard model of particle physics in the limit of flat space time. In practice we would hope to come up with as many such theories as possible, and then we can use them to make predictions which can be compared against future experiment, and thus eliminate the possibilities.

The first way we would naturally attempt to combine the particle physics standard model with general relativity is to write down the standard model in curved space time, and apply Einstein's field equation, and see what we get. The first part of this is straightforward (if we ignore the complications). It is easy to re-construct the Standard Model Lagrangian in curved space time: although it takes a little bit of effort to demonstrate this, the end result is as we would naively expect: all one has to do is multiply each term in the action by the square root of the determinant of the metric (to represent the change in the integration measure d^4x) and convert every derivative operator ∂_μ into a curved space derivative ∇_μ. There are some technical issues relating to the formal construction of quantum field theory in curved space time (in particular the standard construction of the vacuum state assumes a certain set of symmetries which are no longer valid), but these

[10] A very rapid expansion of the universe in its early stages.

[11] A singularity which marks the beginning of space and time.

[12] A non-observable field which carries mass and thus influences gravity but otherwise doesn't seem to interact with other matter.

can be resolved and alternatives put in their place (Wald, 2018). The last of our three requirements is satisfied immediately: flatten space time, and we return to the standard model. Next we can think about the classical limit. The first of Einstein's principles, that as $\hbar \to 0$, objects naturally travel along geodesics, follows from the principle of least action, minimising the action with respect to variations in the fields, just as the laws of motion do in flat space time. So now we need to turn to Einstein's field equation.

We can calculate the stress energy tensor in curved space time in the same way that we would in a flat space time, and we obtain it in terms of the probabilistic quantum fields. And now we come to the problem. In the microscopic theory, the right hand side of Einstein's field equation is built from indeterminate quantum fields. The left hand side, however just depends on the metric and its derivatives, which we are treating as a fixed and thus deterministic background. Saying something indeterminate equals something deterministic doesn't make sense. So this seems to imply that we need to replace the metric with a probabilistic field.

And this is what people have been trying to do. There are several possibilities: either we can treat the metric itself as a quantum field, or the coordinate transformation $\frac{\partial x'^{\mu}}{\partial x^{\nu}}$ (and Einstein's field equation emerges from the principle of least action acting on the gravitational field), we can embed our universe as a surface in a larger dimensional space time (as is done in string theory), or we can introduce some gauge field whose change under a local coordinate transformation cancels out the effects of the Christoffel symbol (just as happens in an electromagnetic gauge transformation). The most straight forward ways suffer from the difficulty that a naive quantum gravity is not renormalisable (violating my demand for mathematical consistency; the number of potential divergences you get from calculating loop diagrams is larger than the number of counter-terms you can add to keep these under control; alternatively the sequence of infinitesimal transformations of the fields doesn't have a unique fixed point), so it is seemingly impossible to reproduce calculable results; but even here, people have been working around this issue. There have been a number of theories suggested, and much progress made in most of them. However, we are no nearer to saying that any of them are right, and I don't think that any of them have yet made any predictions which can in principle be tested experimentally.

I have always been a little bit uneasy about the idea of treating space-time itself as a quantum field. Time can't be the measure of change if it itself is something that can be created or destroyed at each moment. Equally, the metric is just an artefact of our choice of representing physical space and time algebraically. Physical space and time are the real things. The metric is just a tool we have introduced into the mathematics to describe how they are linked together and how to define distances in more than one dimension. But that is just it: it is a tool, a description. Not something (like location or the Fock state for a quark or photon) that is a direct mapping from reality, but something introduced in the process of making a representation to enable us to make calculations. And as such, it is not something that can survive the mapping from the algebraic representation back to reality. Yet those approaches to quantum gravity which look to quantise the metric in effect make it a creation operator for a physical particle, contradicting how it first emerged in physics. All the approaches based on this idea have used the differential of the action with respect to the metric to obtain the stress-energy tensor.

However, we should also consider the other way of deriving the right hand side of Einstein's field equation – from Noether's theorem, by imposing a local symmetry transformation. And in this case, the symmetry makes good theological sense physics is independent of the coordinate system we use. God, classically, is outside of space and time. We can't apply a space time coordinate system to God, and this means that our understanding of God should be independent of whatever coordinate system we use, and this includes curved coordinates as well as Cartesian or hyperbolic coordinate systems. A corollary is that God does not have any preferred coordinate system. This means that the laws of physics, which are

a description of God's sustaining of the universe, should be independent of the coordinate system, which means that the predictions of the quantum field theory, which manifest themselves in the quantum amplitudes associated with a particular action and measure, should be invariant under infinitesimal local coordinate transformations. If we apply Noether's theorem to this rule, we get the stress-energy tensor, proportional to the right hand side of Einstein's field equation. The question is can we obtain the left hand side of the field equation from Noether's theorem?

Of course, I have made a mistake here. The right hand side of Einstein's field equation isn't the stress energy tensor. I have neglected poor old Newton's constant, sitting there innocently in our equations since the early 1700s. I am not alone in neglecting the importance of this number, but perhaps we ought to pay a little more attention to it. What if Newton's constant isn't what it claims to be, namely a constant? What if it in fact emerges from the expectation value of a field? In other words, we can rearrange Einstein's field equation as

$$T_{\mu\nu} + \frac{1}{8\pi G}\left[R_{\mu\rho\nu\sigma}g^{\rho\sigma} - \frac{1}{2}g^{\alpha\beta}g^{\rho\sigma}g_{\mu\nu}R_{\alpha\rho\beta\sigma} + \Lambda g_{\mu\nu}\right] = 0$$

$$T_{\mu\nu} + \frac{1}{8\pi G}\left[R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu}\right] = 0. \tag{687.1}$$

Now instead of converting the metric into a quantum field, why don't we convert the inverse of Newton's constant into a quantum field? This removes the conceptual problem of just putting Einstein's field equation into curved space time - the stress energy tensor is probabilistic, and the gravitational term is also probabilistic. So for the microscopic equation, we replace $\frac{1}{8\pi G}$ with some field $\phi_G^\dagger\phi_G$. What field could this ϕ_G be? None of the standard model fields actually work here: from the dimensions (i.e. counting units of length) we can see that we need a Bosonic field, but putting a gauge Boson in here would break gauge invariance.

That leaves the standard model scalar field, ϕ, and here things get a bit more interesting. The only possible term in the action which could link this field with gravity is

$$\int d^4x \Omega R \mathrm{tr}\phi^\dagger\phi,$$

where R is the same Ricci scalar that plays a role in general relativity, and Ω is some dimensionless constant. This term has actually been hypothesised for some models of cosmic inflation as an addition to the standard Einstein-Hilbert action. What we need to ask is how this term transforms under the same gauge covariant translation that is used to generate the stress energy tensor as the Noether current. What we find is that the Ricci tensor R transforms to

$$R \to R - \nabla^a\epsilon^b(R_{ab} - \frac{1}{2}g_{ab}R),$$

in other words, Einstein's tensor. So the transformation of the action, S, under this symmetry is

$$S \to S - \int d^4x\sqrt{-\det g}(\nabla^a\epsilon^b)\left[\Omega\mathrm{tr}\phi^\dagger\phi(R_{ab} - \frac{1}{2}g_{ab}R) + T_{ab}\right]. \tag{687.2}$$

Now I want to restrict the action to those which remain unchanged under any local change of the coordinate system. God is outside space and time, and so will react with nature in the same way regardless of choice of coordinate system. The action describes how God sustains nature, so must be invariant of the choice of coordinate system. So, one might argue whatever the value of $\nabla^a\epsilon^b$ which parametrises the change in the coordinate system, the change in the action must be zero. And this means that

$$\left[\Omega\mathrm{tr}\phi^\dagger\phi(R_{ab} - \frac{1}{2}g_{ab}R) + T_{ab}\right].$$

Take the expectation value to get the classical limit, rearrange things a bit, and choose the right value of the parameter Ω, and you get Einstein's field equation and thus a quantum theory of gravity. All nice and simple.

Unfortunately, it is not quite that simple. The problem is, to be consistent with experimental data for cosmic inflation, the parameter Ω needs to be within a certain range. To be consistent with the measured value of Newton's constant, the value of $\langle \Omega \mathrm{tr}\phi^\dagger\phi\rangle$ needs to be just right. Thirdly, the observed Higgs mass is related to the value of $\langle \mathrm{tr}\phi^\dagger\phi\rangle$. So, if this model were correct, we would be able to calculate the value of the Higgs mass from the strength of gravity and the cosmological data. And we would get it wrong. By about a factor of 10^{-17}.

So we have to start again, but I think that this calculation is along the right lines. It is just not the standard model scalar field that drives gravity, but a different scalar particle. I have chosen to express this field as φ. Unlike the standard model field, which has an SU(2) (spherical) gauge symmetry, φ has just a U(1) (circular) gauge symmetry. This means that it cannot interact with any of the standard model matter fields except (possibly) the standard model scalar field. We could introduce a coupling between this field and the Fermion fields, but that would break the gauge symmetry since it transforms under a different gauge group. We could make the photon the U(1) gauge field associated with φ, but that would also cause issues (not least by generating a rather large photon mass in the same way that the Higgs mechanism generates a mass for the W and Z Bosons). So in addition to the new scalar field, we also have to introduce a new U(1) gauge field D_μ.

But this is then, I think, what we need. Such a field is not forbidden by any symmetries, so there is a case that it should be there. So, if S_{SM} is the standard model action, then I propose that the action that describes reality is

$$S = S_{SM} + \int d^4x\sqrt{-g}\Big[\Omega R\mathrm{tr}\phi^\dagger\phi + \frac{1}{4\pi g'^2}\mathrm{tr}(\nabla_\mu D_n u - \nabla_n u D_\mu)^2 \\ \alpha R\varphi^\dagger\varphi + \frac{1}{2}\varphi^\dagger(\nabla_\mu - i\alpha_D D_\mu)^2\varphi - \frac{\omega^2\mu^2}{2}\varphi^\dagger\varphi + \frac{\rho}{4}\varphi^4 + \beta\varphi^\dagger\varphi\mathrm{tr}\phi^\dagger\phi\Big], \quad (688.1)$$

where Ω, α, β, ρ, g' and ω are additional dimensionless constants to add to the collection that we already have.

This would lead to a change in the action,

$$S \to S - \int d^4x\sqrt{-\det g}(\nabla^a\epsilon^b)\Big[(\Omega\mathrm{tr}\phi^\dagger\phi + \alpha\varphi^\dagger\varphi)(R_{ab} - \frac{1}{2}g_{ab}R) + T_{ab}\Big], \quad (688.2)$$

which leaves us where we were before, but this time we can use the φ field to drive gravity, keep Ω small enough that we can have the measured Higgs mass, and generally satisfy the experimental parameters.

One final technicality is that it is not so much the action that we need to be invariant under the symmetry but the amplitudes describing the likelihood that we get from an initial state to a final state. They are, after all, the physical things (what we can measure), while the action is just a mathematical convenience (a means to calculate the things we can measure). The process of making the amplitudes invariant leads to a expression known as a Ward Identity. At the heart of the Ward Identity is the change of the action, so it leads nearly the same result as action invariance, but with a small correction. That correction in the case of this symmetry is a term proportional to the metric, in other words a term leading to a small cosmological constant.

What this means in practice is that we are further restricted in the possible choices of coordinate system used to represent space-time algebraically. We already know that we need to choose the coordinate system to reflect the dimensionality and topology of physical space time. In addition, we are restricted by the need to respect general coordinate invariance and the quantum version of Einstein's field equation to ensure that the (still with a great

deal of freedom) choice of coordinate system doesn't leak through to affect the results when our calculations are mapped back to make predictions in the physical world. Matter plays a role here. In each possible path from initial to final state, there will be a slightly different distribution of matter, and thus a slightly different set of metrics and coordinate systems will be allowed if we are to respect the symmetry. We know that one path will happen in practice, and thus one set of metrics is allowed. We just can't predict in advance which one.

The coupling between the curvature and the scalar fields acts similarly to an additional mass term for the scalar field, in other words it will modify the observed mass of the scalar field. When the curvature is relatively small (as it is today with gravity being weak), the effect of this on the masses of the scalar fields is not really measurable. But when you go to large gravitational fields, as you get around black holes or in the early universe, there is a large change in the scalar field mass. This means that there is a large change in the value of $\varphi^\dagger\varphi$, and thus in Newton's constant. What this means is that Newton's "constant" could have been much smaller in the early universe when matter was tightly bunched together and the curvature of the universe bigger. Depending on the values of the constants, gravity could even have been repulsive rather than attractive in the early universe. This implies cosmic inflation.

The new φ and gauge fields are candidates for the "dark matter" required by the various studies of galactic dynamics and stability.

Thus this model has the potential to explain all the known experimental results. It reduces to the standard model in the limit of flat space time, general relativity in the classical limit, and includes a model to explain dark matter, dark energy and cosmic inflation.

In other words, by assuming one additional symmetry which makes sense given our presumptions and adding an allowed additional field into the action, we have derived Einstein's field equation from an extended standard model. The local symmetry is one which is fully expected from the theology. In short, if this idea is correct, we have completed our derivation of modern physics from premises derived from classical theism, and thus shown that classical theism, and its picture of physics as a description of how God sustains the universe, is in no way inconsistent with modern science. Is this idea correct? That will take a lot more work to demonstrate: it is always easier to write a theory down than prove it right. Firstly, it needs to be tested against what is already known: does it mess up the current very good predictions of the standard model? Then we need to devise predictions which can be tested experimentally. And I have not even begun to address whether or not this theory is renormalisable and mathematically self-consistent.

However, if this idea works out, then we have explained Einstein's theory of gravity in terms of Aristotelian, rather than mechanist, metaphysics. Even if it doesn't work out, and we need to quantise the metric, it is still likely that that will fit in better with Aristotle than mechanism. It is very likely that the final theory will be a quantum field theory (central to particle physics), and the symmetry of general covariance (central to general relativity) will be heavily involved. But we will need to know the exact theory to work out the details. Einstein's gravity, of course, is usually thought of as the natural culmination of Galileo and Descartes' Metaphysics. If, on the other hand, it has been shown to ultimately originate from premises which are very much contrary to the philosophical movement started by Descartes, then we have to question whether the mechanistic philosophy is indeed demanded by the success of science.

15.10. Recap

The God of classical theism has a number of essential attributes:

(1) He has a free will.
(2) He sustains the universe, so that beings only continue to exist as long as he chooses them to exist, and all material objects are subject to God's will.

(3) God can create and destroy matter as He pleases, subject only to the need for self-consistency and consistency with God's nature.
(4) He is rational.
(5) He exists outside of time and space, and is unchanging, and therefore all material beings will perceive Him in the same way (all other things being equal) regardless of when they are, where they are, how fast they are moving or accelerating, or any other properties.

Material objects, however, are changeable, and can exist in a number of different physical states.

Because of God's free will and omnipotence, we cannot, even with a complete knowledge of the universe at a given moment in time, predict with certainty what the universe will be at the next moment in time. If God wants to move an electron from a location A to a location B, he can do so via any route that he pleases, regardless of whatever is happening elsewhere in the universe, or what matter is around. This is contrary to the situation in mechanistic physics, where the path is fixed and determined by the equations of motion.

However, although we cannot predict the way God evolves the universe with certainty, we can still express our uncertainty mathematically, either as a probability or (more accurately) a probability amplitude or likelihood. A likelihood is a measurement of uncertainty, which, although it itself does not satisfy the strict mathematical definition of probability, is still related to the probability and probabilities can be calculated from it. It is more useful in modelling physical systems. We can therefore express our knowledge of the state of the universe as a probability amplitude, which is a list of all the possible states the universe could be in together with the likelihood that it is in one of those states.

Observed properties can be computed from the knowledge of the physical states, but not as in classical mechanics where all the properties of a being can be defined for each of the states. We have considerable freedom in how we represent the probability amplitudes, and in some of those representations certain observable properties will be well defined but not others, while in a different representation other properties of the being will be well defined. To measure a particular property, we have to first of all switch to a representation of matter where that property can take a definite value. It is therefore a serious mistake to define a being in terms of a list of properties.

If the universe starts in one state in one moment of time, God can evolve it to any other state at the next moment of time. There is therefore some likelihood that at the next moment in time the universe will exist in any particular physical state. God will move it to one particular state, but since we don't know what state this is, we have to consider all possibilities. The likelihood that the universe will be in a particular state thus depends only on the relationship between God and matter, and the initial state of the universe. Thus we can develop a set of rules allowing us to says things like, 'If that is the initial state of the universe, then this is the likelihood that it will be in a particular final state after God has finished with it.' This implies that there is an abstract operation (the time-evolution operation) which we can construct which allows us to start with any state in the universe, and to compute the likelihood that the universe will finish in any other state. The nature of this operation can only in principle depend on the initial state of the universe, the relationship between matter and God, and the nature of God. However, because this operation is general enough to cover every possible initial state of the universe, the form of the operation itself will not need to specify which initial state it is applied to. Since God is immutable, God's nature, and therefore His side of the relationship between God and matter, will not change based on the state of the material universe. The evolution of matter is driven by God's will alone. Therefore, the possible form of the time-evolution operator can be restricted to a small number of possibilities by considering the attributes of God. In particular (and I should emphasise that this list is only a summary of the argument of the previous sections, and contains a number of simplifications and skipped details),

(1) The universe must start in some definite initial state and evolve into some definite final state. This means that the operation which describes the likelihood of God evolving the universe preserves the total probability amplitude (the sum of all probabilities must remain one), which restricts it to the class of operators known as unitary operators.

(2) Because God is outside of space and time, and therefore to God one moment in space and time is the same as any other, the time-evolution operation is the same for every moment of space and time, which means that it can be expressed in terms of a sum over all possible locations and times of operators representing the possible ways in which God can create and destroy matter.

(3) Because God is outside of space and time, the time evolution operation cannot depend on an absolute location or absolute time. Instead, it can only be a function of relative displacements between particles. Since every relative displacement (given certain conditions which are satisfied in practice) can be expressed in terms of local differential operators, this means that each term in the time-evolution operation can be expressed so that it only contain interactions between matter at one particular location and the same time. Mathematically, this principle of locality demands the conservation of energy and momentum, and consequently the principles of efficient and final causality.

(4) Because God is outside of space and time, matter relates to God in the same way regardless of its velocity or location in the universe, and thus the time-evolution operation is the same whatever coordinate system we use (a coordinate system describes how a particular observer perceives the rest of reality), and does not change when we shift from one coordinate system to another. This applies to both global and local coordinate transformations.
   - Global coordinate transformations of the universe are where we change the coordinate system in the same way everywhere. Demanding that the time evolution operation is of the same form after a global transformation leads to the theory of special relativity.
   - Local coordinate transformations are where we change the coordinate system in different ways in different locations. Demanding that the time evolution operation is of the same form after a local transformation leads to the principles of general relativity and the theory of gravity.

(5) The individual states of the universe don't just depend on location and time, but also other quantities. Some of these quantities can be mapped to continuous variables with a circular or spherical topology; i.e. each individual state can be represented by one particular point on the circumference of a circle. Just as we can use different coordinate systems to map from physical space-time to our representation of it, so we can use different representations to map physical states to different points on the circle. So in one representation the top of the circle will represent a state A, and in another representation the bottom of the circle will represent that state. There is no need to use the same representation of these circular or spherical properties for every different location in space and time state. Because God is outside space, time and the space that describes these new circular properties, matter in any representation of these properties will relate to God in the same way, which means that the time-evolution operator must be identical regardless of which representation we use. This includes local (dependent on space) transformations. The only way to reconcile this with what we have developed so far is to introduce a gauge principle, and new particles, the gauge Bosons, which interact with the matter fields.

(6) The time-evolution of the universe can thus be reduced to the possibility that God can, at each location in space and each moment in time, destroy matter and gauge

Boson particles and replace them with new matter and gauge Boson particles, subject to the conservation of energy and momentum.

(7) Because God is outside of space and time, He has no preferred distance scale. Therefore the states which describe physical matter and the time evolution operator must be expressible in a way that is independent of any intrinsic distance scale. The intrinsic distance scale in a field theory is known as the correlation length (the name is taken from an analogy with statistical physics, and is slightly misleading in its applications in particle physics), and the state of the universe independent of scale has a infinite correlation length. The representation of matter with infinite correlation length can be found by:
   - Start with a self-consistent representation with finite correlation length.
   - Construct a self-consistent transformation of the representation which increases the correlation length.
   - Apply this transformation an infinite number of times, until an equilibrium is reached.

   The infinite number of applications might seem daunting, but in practice it approaches a well-defined mathematical limit which can be computed or approximated from knowledge of the operation. This process is known as renormalization, and it homes in onto one particular representation of the possible physical states, which is the only one which could correspond to the physical universe.

(8) By considering every possible way in which God could evolve the universe, and the likelihood for every possible final state, we can compute the probability that the universe can finish in a given final state.

(9) There are certain paths which God could use to move the electron from A to B which are known as the minimum action paths. In the sum over likelihoods, the paths which deviate considerably from the minimum action paths cancel each other out, meaning that the dominant contribution are those paths which are close to the minimum action paths. These paths are those predicted by classical physics. If our experiment only has a large resolutions for location, momentum, time and Energy then the results of our measurement will seem to imply that all particles travel by the classical path (consistent with mechanistic physics). However, if we look at a finer resolution, we see that in reality the universe is indeterminate, unpredictable, and there is no single path which matter takes to get from A to B.

By using these principles, we can restrict the physical theories consistent with classical theism to just a few possibilities. Add in a few additional assumptions drawn from observation:

- that we live in three space and one time dimension;
- that there is a $U(1) \otimes SU(2) \otimes SU(3)$ gauge symmetry (which pictorially is represented by combination of a circle, whose circumference is one dimensional, a hyper-sphere with a three dimensional surface, and a hyper-sphere with an eight dimensional surface);
- that there are three families of Fermionic matter and one scalar field,

and we arrive with the mathematical form of the standard model of particle physics combined with general relativity. Both theories have been tested to extra-ordinary experimental precision. The only limitation is that there are a number of parameters (a little over twenty) which describe the strength of the interactions between different types of particles and cannot be computed directly from metaphysical principles. However, baring this caveat, what I have shown is that if classical theism is true, combined with a few additional assumptions taken from the most basic observations, then we can expect the universe to be precisely what we observe it to be. If classical theism is false, then it is unlikely that the alternative metaphysical picture would imply the correct physical theory. The indeterminacy of matter implies that the universe is guided by a free and intellectual will, which seems to rule out atheism and deism; that the same rules apply to all matter seems to rule out polytheism;

that the symmetries used to constrain the physical theory were built on the idea that the divine will is unchanging and outside space and time seems to rule out more modern forms of theism. We seem to be left with classical theism as the only viable option.

Of course, 'seem' is not a proof. There might be (and are) alternatives which are both self-consistent and consistent with modern physics. It may be that there is a mechanistic explanation for physics; while the naive mechanism that was accepted until the early twentieth century fails, there might be other forms of mechanism which are consistent with quantum physics (the Everett scenario, which denies that the multiverse is indeterminate, is the only possibility I know of – but I make no claims to omniscience, so what is unknown to me might be known to others); it is difficult to rule out unknown alternative explanations if your only tool is falsification. All I have done is to demonstrate that an Aristotelian classical theism seems to work well, whether it is induced from physics, or we use it as a premise to construct physics. What we can, and should say, is that there is no longer reason to believe that physics prefers mechanism (which underlies most modern forms of atheism) over classical theism (which underlies many forms of Christianity). My inclination is once we know the full truth of physics (and we are still far from being there), mechanism (and thus most of modern atheism) will either be found to be untenable, or at least will require heavy modification to maintain consistency with physical theory. There is no good reason based on the form of the laws of nature to favour enlightenment philosophy over branches of the older Aristotelian tradition. There are good reasons to favour the Aristotelian tradition.

What of the other reasons for favouring one metaphysical construct over the other? Obviously, first of all, we need to turn to philosophy itself, and here my own studies and reading suggest that the Aristotelians have a clear advantage. There are numerous issues in modern philosophy – free will, in the philosophy of mind, epistemology, ethics, the relationship between substances and properties and so on – which are absent in the older system. On the other hand, I am not aware of any internal contradictions or tensions in the mature forms of Aristotelian philosophy (the issues I have seen raised by modern philosophers tend to be either straw men, based on misunderstandings, or interpret classical philosophy by incorporating assumptions from modern philosophy). But this discussion is best done by the professionals rather than amateur philosophers such as myself.

But there are other arguments which also need to be considered, which stray more towards the physics side of the philosophy of physics. First, I wish to mention the anthropic principle.

15.11. Fine Tuning

What are the chances of there being new fundamental particles beyond the Higgs (aside from whatever is needed to explain gravity)? Here lies one of the major remaining puzzles of particle physics; it is known as the hierarchy problem. The standard model has a natural mass scale, as we have seen – related to the Higgs mass. Cosmology also has a natural mass scale, ultimately derived from Newton's gravitational constant, and known as the Planck mass; and this is believed to be (approximately) the energy scale where gravity will start to interfere with our particle physics experiments. The first difficulty is that the observed Higgs mass is approximately seventeen orders of magnitude smaller than the Planck Mass (the precise figure depending on how the Planck mass is defined). This is a big number (roughly the same as the difference between a hand-span and the typical distance between neighbouring stars), and it does not seem natural that the zoo of fundamental particles should be empty for such a large scale until we come to whatever mediates Gravity. The second part of the problem is that the measured Higgs mass is not the same as the figure we put into the theory. Instead, the measured value receives corrections from both renormalization and the Higgs self-energy (which describes diagrams where you both start and end with a single Higgs, but something interesting happens in the middle, for example, the Higgs decays into a top/anti-top pair, and these then recombine to give another Higgs

Boson). The self-energy gives another contribution to the propagator which resembles the square mass in a theory, and the square of the measured mass will be the sum of the square of the bare mass (i.e. $2\mu^2$) of the Higgs and the self-energy. The amplitude for these self-energy contributions from other particles depends on the coupling between the Higgs and those particles - which is proportional to the particle mass. Thus the heavier the intermediate particle is, the more it will contribute to the Higgs self-energy, and the more it will push up the observed Higgs mass. Thus we expect the measured mass of the Higgs to be roughly the same size as the mass of the heaviest particle in the system; in the standard model that particle is the top quark whose mass is a bit larger than the measured Higgs mass; the top is a little bit too heavy to make physicists comfortable (it is on the borderline of what is allowed for the Higgs vacuum to be stable), but not by so much as to cause major distress (it might, once everything is calculated properly and gravity included, be just the right side of the border rather than just the wrong side).

If there is some additional particle which mediates gravity, having a mass of the order of the Planck scale which it obtains by a coupling with the Higgs Boson, then we clearly have a problem because we would then naturally expect this contribution to the Higgs self energy to drive the measured Higgs mass to the order of the Planck scale, which it isn't.

The contributions from the self-energy are problematic. Firstly, the integrals we find are highly divergent they scale as the square of our cut-off energy – and the most natural cut-off energy to apply if there is no physics except for the standard model and gravity is the Planck Mass. This can in principle be solved by renormalization, but requires the bare mass of the Higgs to be tuned to an incredible degree (to one part in around 10^{16} in order to get the measured value). Secondly, as discussed, the finite part of the self-energy - what is left over after renormalization – leads to a correction to the Higgs mass proportional to the particle mass. This means that if there are particles with larger mass than we know about, they would (baring a seemingly astounding cancellation) lead to a much larger Higgs mass than we observe. Now one important point is that the most dominant contribution of Bosons to the Higgs self energy can have either a positive or negative sign, so in principle it is possible to cancel out a Fermionic contribution to the self-energy by a Bosonic contribution to the self energy (the Higgs decays into a number of Bosonic fields, which then recombine to make another Higgs).

So we seem to have only a few options: 1) the standard model is complete except for gravity, and the bare Higgs mass is tuned to an incredible degree; 2) the standard model is not complete, with both Bosons and Fermions as yet undiscovered, and the masses of the additional particles are tuned to an incredible degree so their contributions to the Higgs self-energy very nearly cancel each other out; 3) the standard model is not complete, and the parameters of the remaining particles are linked to the parameters of the particles we know about by some symmetry in order to achieve the near-total calculation; 4) Some ideas which consider different numbers of spatial dimensions solve the hierarchy problem; 5) We misunderstand the physics. Probably the idea that has received most attention is that there is some symmetry (known as supersymmetry) linking the Bosonic and Fermionic contributions to the Higgs self energy. The most natural forms of this idea should be either confirmed or falsified experimentally in the next few years; so far a few of the simpler models have been ruled out[13]. I think that option 5) is also a possibility. The worst case scenario was based on the idea that the effects of quantum gravity enter at the Planck scale, but

[13]Supersymmetry doesn't wholly solve the hierarchy problem, though, even if it is found to be correct. Supersymmetry must be broken, otherwise we would have observed all the hypothesised partners already, and the masses of the partners will be around the supersymmetry breaking scale. Above this scale, there would indeed be cancellations in the radiative effects to the Higgs self-energy between the observed particles and the hypothetical super-particles; below this energy scale we would expect the corrections to the Higgs mass to appear. Thus the natural scale of the Higgs mass would be the supersymmetry scale. Thus supersymmetry converts one problem into two: a) Why is the Higgs so much lighter than the not yet observed supersymmetric partners? and b) Why is the supersymmetric breaking scale so much lighter than

this scale just comes from throwing a few numbers together, and not any real theoretical motivation. It may be, for example, that whatever mediates gravity doesn't couple with the standard model Higgs. There are also issues concerning the scale (i.e. renormalization group cut-off) dependence of the Higgs couplings λ and μ which I am not convinced have been properly considered (Altmannshofer et al., 2015). Rather than speculate about new physics, I will assume that the standard model is complete baring some term to explain the neutrino masses, and some term to describe Gravity. That means that we might well need fine tuning of the Higgs mass.

So far, I have just shown that the general form of the theory is in line what we expect from classical theism. But my argument was not demonstrative: I just showed that the standard model Lagrangian was consistent with the underlying principles, but I did not (and could not) explain why it took the precise form that it did. At various points in the argument, I noted that something had to be either this or that, and I selected one of those choices, chiefly because I knew that led to the right answer. I have shown the consistency of physical theory with premises consistent with classical theism, but not deduced the exact form of it from those premises.

Now, it should noted that this is not just an issue for classical theism, but for any underlying metaphysics. The best that a metaphysics can do is provide various axioms, such as symmetry requirements or an overall framework, such as either the path integral approach that follows from the theistic philosophy I propose here or the differential equations of motion and forces that arise from a mechanistic philosophy, and there will undoubtedly be many possibilities consistent with those requirements. For example, suppose that we had a theory consistent with the symmetry requirements which contained ten different types of fields. What if we removed one of those fields, either by setting all the relevant coupling constants describing the interactions with other particles to zero, or by just not writing them in the equations? For many of them (one exception is if we tried to remove gauge fields), we would still have a theory consistent with the symmetry requirements, it would just be different from the original. The best we can say is that certain types of matter can be there, but not that they actually are there; and I do not believe that metaphysics, whether atheistic or theistic, can answer this question without bringing in some additional assumptions.

For example: why three spatial dimensions and one time dimension? Why a $U(1) \otimes SU(2) \otimes SU(3)$ gauge theory? Why left and right handed Fermion fields? Why three families of Fermions (i.e. up, down quarks, electron and electron neutrino make one family, the strange, charm, muon and muon-neutrino the next)? Why is the sign in front of μ^2 in the expression for $\hat{\mathcal{L}}_{\text{Higgs}}$ negative rather than positive in equation (674.4) (it makes a huge difference to the physical consequences of the theory)? Why this particular pattern of spontaneous symmetry breaking (particularly important in the Higgs field and origin of mass, but also elsewhere in particle physics)? Why only one set of standard model particles (i.e. there could be another set of electrons, quarks, Higgs and so on with no couplings to our own standard model except through gravity) – if, indeed, there is only one? And what about all these constants – the gauge couplings α_W and so on, the couplings between the Fermion fields and the Higgs field? So far in our analysis, these can take any values the only one we don't have to to tune is the scale setting parameter μ, the mass of the Higgs Boson, which essentially plays the role of our ruler, establishing the natural units to measure distance.

So we have thirty two numbers from particle physics and cosmology – two coupling constants α and α_S[14] between the particles and the gauge fields; θ_W; the Higgs interaction

---

the Planck scale associated with gravity? In light of this second question, supersymmetry doesn't so much solve the hierarchy problem, but simply rephrases it by creating a new problem in its place.

[14]In practice, these depend on the energy scale used in renormalization; so we have to suggest some fixed energy scale to define them - the Higgs mass; the mass of the neutral weak interaction gauge Boson; or zero are all natural choices which have been made. But ultimately, it doesn't matter what we choose:

parameter λ; three electron and six quark couplings to the Higgs field; three additional parameters to account for the neutrino masses; four parameters for the CKM matrix; the θ_2 coupling for the possible $\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma}$ term in QCD (which seems to be 0); in the model for gravity proposed in section 15.9 there are five more parameters describing the interactions of the new fields; and from cosmology the physical baryon density; the physical dark matter density; the dark energy density; the scalar spectral index; the curvature fluctuation amplitude; and the re-ionisation optical depth. All of these parameters could in principle (as far as we know) take any value between zero and something very large and possibly infinite, and we would still be left with a self-consistent theory. There may well be links between these parameters; for example some of the cosmological parameters might ultimately be found to depend on the standard model parameters (or the parameters of some extended standard model). Equally, of course, the various models might well not be complete (indeed, most people don't believe that they are complete), which might lead to a few more dimensionless parameters being required in the full theory. But, at our current level of knowledge, our understanding of physics boils down to these thirty-two numbers. None of them are predicted from theory. Most of them can't be predicted from physical theory even in principle.

Change any one of these parameters by even a small amount and the nature of the observed universe changes dramatically. Of course, we cannot run experiments to demonstrate this – in experiment, these parameters are fixed and we are stuck with them. But, just as we calculate observables theoretically with the observed parameters to get a good match with experiment, we can just as easily perform calculations with different values for these parameters and see what the universe would look like in that case. But how different would the universe be? The answer is almost certainly one in which any life would be impossible.

For example, let us consider the requirements for heavy elements to form in stars:

(1) The process of nuclear fusion requires a careful balance between the electromagnetic, strong and weak forces. In a star, protons and neutrons rush around at high energies. We start with hydrogen, the simplest element. To get a new element, Helium, we need to add a proton to this. The way this works is that first of all a neutron bangs into a proton, and gets bound by the strong nuclear force, to create a deuteron nucleus, and then two of these combine to form a stable Helium nucleus. The protons have a large electrical repulsion, and a much shorter ranged attraction due to the strong nuclear force. They have to have enough kinetic energy to overcome the electric repulsion between them, so that the nuclear force can bind them together, but not too much kinetic energy that they can also break apart the strong interaction string. If the strong nuclear force were a little bit stronger in comparison to the electromagnetic force, then the process of nuclear fusion would occur too quickly, and stars would run out of fuel too soon before life could evolve; if it were a little bit weaker, then the process of nuclear fusion would be too weak, and stars would not be able to ignite, creating the high kinetic energies which allow the process to continue.

(2) Then we have the contribution of the weak nuclear force: we need two neutrons to form a stable helium nucleus. These come from the decay of protons into neutrons via the weak nuclear force. This very much controls the rate at which fusion can occur. If the weak interaction is too strong, then again, the stars burn up too quickly; too weak and nuclear fusion can't really get going because there are too few neutrons around to merge in the process.

These physical processes are dependent on six of our twenty-nine fundamental parameters: the strengths of the strong, weak and electromagnetic forces, and the couplings of the up and down quarks and electrons to the H Boson. We have two independent conditions each

how the coupling constants change with this scale is known, and given them at one energy scale, we can in principle calculate them at any other. The equations describing the evolution of these constants each have one dimensionless free parameter – and it is ultimately this number which I am referring to.

constraining these six parameters to within a narrow band. If the first of these conditions is broken, then there will be nothing in the universe except hydrogen. If the second of these conditions is broken, then there will be nothing in the universe except hydrogen and helium. The first of these conditions means that if we want to have a more interesting universe than a gas of hydrogen and helium, we have not got six free parameters, but only five. In other words, fix five of these numbers, and the sixth is determined within some rather narrow bounds. When we add in the second condition, we are left with only four free parameters – we can fix four of them to whatever we like, and then the last two have to take specific values (at least within a narrow range). Note that there is no fundamental physical or metaphysical bound – this bound is only if we add in the assumption that we want to be able to form more complex chemicals than just hydrogen and helium. But there is more.

(3) The next part of the chain of the generation of elements is the formation of carbon. Three helium atoms combine together to make a carbon atom. The chances of three atoms coming together in the same place within a short enough period of time to do this is so remote that we can forget about it, unless there is some other factor in play. In practice, there is some other factor in play. In nuclear and particle physics, the probability of an interaction occurring is given in terms of what is known as the cross section. This cross section is a function of the energy of the incoming particles. It is not, however, a smooth function, but has certain distinct peaks corresponding to the energy levels of the Hamiltonian associated with the final state. In other words, if you happen to have (near enough) just the right energy to drop into an eigenstate of the Hamiltonian, then there is a much higher change of creating the particle than if you tried a different energy (where you would have to form a linear supposition of eigenstates, which isn't stable). Now it happens that carbon has an energy level (an excited state) at just the right energy which corresponds to the mass of three helium nuclei (recall that mass is a form of energy). This is fortunate, because without it would not be possible to form carbon, and we would not be here. So why does carbon have this Hoyle state? Well, ultimately the energy levels of carbon are a complicated function of the strong, weak and electromagnetic forces, together with the lightest quark and electron masses. So to say that carbon must have a resonance at just the right point means ultimately that these six numbers must be related to each other in just the right way.

(4) It happens that the mass of the proton is slightly lower than the mass of the neutron. Almost all of the mass comes from the binding energy due to the strong interaction. The small difference comes mainly from the slight mass differences of the quarks, and electromagnetic effects. Change these numbers a little bit, and we can change the mass difference between the proton and the neutron. Make the neutron a little bit lighter than the proton, and the proton would not be stable (since it can easily decay into the neutron), meaning no hydrogen. Other elements might still exist, but no hydrogen means no water, and no amino acids, no DNA, and so on. On the other hand, make the proton too much lighter than the neutron, and the di-proton state (with two protons) becomes stable (although this also depends on the strength of the fundamental forces), and again stars would burn up their supply of hydrogen too quickly. Again, the mass difference has to be just right, and this puts a tight bound on the six parameters.

(5) On a related note, neutron decay also requires the mass of the electron to be just right: if the electron were a bit heavier, or the neutron a bit lighter, then the neutron would not be able to decay freely (which requires the production of an electron). On the other hand, make the difference too much, and neutrons would decay too fast. In either case, the fine balance which powers stars would again be broken.

(6) A further coincidence led by the weak interaction occurred in the early universe. We need the right ratios of protons and neutrons in the early universe to have the necessary ratios of hydrogen and helium we observe today. The neutron is unstable, except when bound with the proton to form a deuteron. It took about a hundred seconds after the big bang for the universe to cool enough for the deuteron to form; the half life of the neutron is about 615 seconds. Make the weak interaction a bit stronger, and many more of the neutrons would have decayed before they could have formed deuterons: there would not have been enough heavy hydrogen around when stars were formed to start the fusion process. On the other hand, make the weak interaction weaker, and a different problem emerges: neutrinos are required to re-distribute the energy between the protons and neutrons in the early universe, so that the number of protons and neutrons remain in thermal equilibrium (the weak interaction, which involves the neutrino, is the only way to convert a proton or neutron into each other). To maintain a thermal equilibrium between protons and neutrons requires that they have some interaction. At a large temperature, the small mass difference between the proton and the neutron doesn't really matter, and you will get roughly the same numbers of each type of particle. As the temperature drops, the thermal equilibrium state will increasingly favour having more protons than neutrons. However, to maintain thermal equilibrium between the numbers of protons and neutrons, we need a large density of neutrinos, so they can easily change from one state to another. As the universe expands, the density of neutrinos necessarily decreases. At some point, it becomes too small to convert a proton into a neutron, and the proton and neutron would no longer be in thermal equilibrium: at this point, the ratio of protons to neutrons would remain fixed, until we reach long enough times for neutron decay to become important. If this point when the neutrino density became too low happened too soon (which would involve a weaker weak nuclear force), there would be almost as many neutrons as protons in the earliest stages of the universe. These would join together to form deuterons, and then fuse to form helium. There would be much less hydrogen in the universe, which again means no water, or organic chemicals.

We now have six independent conditions narrowly bounding the parameters from above and below, and six free parameters under consideration (with the remaining parameters playing no significant role in these processes). What this means is that to have a universe containing both hydrogen and carbon (and elements heavier than carbon), these six numbers are constrained to within a set of very narrow bounds. We no longer have any choice: these values have to take the numbers they do (with a very small range of error) if we are going to have carbon based life. With only one or two conditions, fine: we could modify one of the parameters a little, and also change the others, and everything would still be OK. The parameter space would still be constrained, but not so much that it makes it so unlikely that they would all just happen to be right. But we now have six conditions and six numbers, which means that there is no freedom at all: the fundamental physical parameters have to take just the right values for there to be elements made of carbon and hydrogen. These compounds, of course, include the fundamental building blocks of life as it exists on our planet.

Do you want more?

(7) Consider the ratio of the electron and nucleon masses. This helps determine the allowed energy levels of atoms, which again is important: make the isolated atom too tightly bound, and molecules can't form, so there would be no chemistry. Make it too lightly bound, and atoms would become unstable. Again, vary the ratio of the electron and nucleon masses just a little bit either way, and there could be no life.

So we now have seven constraints applied to six parameters. The system is over-constrained. This means that we are very fortunate to be here, since it is not logically guaranteed that there would be a window in which life is possible.

(8) Next, supernova core collapse. Once we have formed the basic elements in a star, we still need to distribute them out. This is done when a star explodes. As the star runs out of material to fuse, the kinetic energy of the particles is no longer enough to beat gravity, and the core of the star (which is by now mostly made up of the heavier elements) starts to collapse under the combination of pressure and gravity, to form a neutron star or perhaps a black hole. As it does so, the outer material of the star also moves in, then rebounds dramatically to create an explosion. This rebounding is caused mainly by the pressure of neutrinos, which emerge at a critical point of the process as the protons convert into neutrons to form a neutron star. Again, the weak interaction plays a key role in this process: too strong and the neutrinos would not be able to escape from the now very dense stellar core; too weak and they would not be able to interact with the outer stellar material, to push it outwards and cause the explosion.

We now have eight constraints for seven parameters (as we now have to include gravity as well, as the strength of the gravitational force affects this last process).

I could add more, but I have sufficiently made my point. The cosmological anthropic coincidences, are, if anything, far more restrictive than the conditions on the standard model parameters I have been discussing. There are over a hundred[15] conditions constraining the approximately thirty fundamental parameters which need to be satisfied if we want to have life (and the constraints are strongest on the seven key parameters I have already discussed).

Why is this relevant? After all, we know that we exist, so the universe has to be just right for us. Are we just like a puddle, sitting in a hole, and reasoning that the hole was made perfectly to fit the puddle, rather than (as in reality) the puddle was made to perfectly fit the hole? An alternative way of phrasing this is: OK, the universe is tuned for our sort of life, but why should we think that our sort of life is the only one possible? Could not something else exist if the universe was just a little different? The problem of this argument is that it fails to consider the nature of the universe if the coincidences fail to hold. If you can come up with a way to form life out of the basic standard model building blocks with only hydrogen in the universe, or only helium, or no molecules being able to form, then good luck to you.

Life requires three things: 1) exceptionally complex information needs to be stored; 2) that information needs to be replicated and copied; and 3) the information needs to be read and 'understood' by physical processes. The storage of information requires some physical means to both do so and to read that information. A long chain of letters is meaningless (even if they happen to be precisely the same order as found in *Hamlet*) unless there is some method of extracting meaning from those letters; of interpreting the code so it has physical or intellectual significance. We can only read *Hamlet* because the words in it have some pre-assigned meaning. For us, that information is stored on the DNA and RNA molecules, interpreted using the interactions between those molecules and the various amino acid chains that build up our bodies, and duplicated through various chemical interactions. The information storage mechanism needs to be sufficiently extensible that there can be a wide variety of life; and it needs to be possible for it to be built up from the fundamental Fermions and Bosons which make up the universe. I see no reason why the specific molecules of DNA and RNA have to be the only ones which enable life to occur; but whatever other

[15]The literature on the anthropic principle is vast, and I cannot list every work of importance. Examples include (Carter, 1974; Barrow and Tipler, 1986; M. and A., 2007; J. and J., 1979; Davies, 2006; Rees, 1999; Bostrom, 2010; Leslie, 2002; Barnes et al., 2016); and many others. The one attempt to argue against the existence or significance of these constraints using standard physics I am aware of is (Stenger, 2011); however this work is weak and makes numerous errors, as discussed in (Barnes, 2012).

possibilities there are would require a similar degree of complexity so they can encode all the information required to build a living organism. A universe containing only Hydrogen, for example, simply cannot support life, since the only thing you can build out of hydrogen is a hydrogen molecule, which contains no meaningful information nor physical means of interpreting that information. Every hydrogen molecule is the same as any other hydrogen molecule. Adding in Helium doesn't help us, since it is chemically inert.

Yes, we can be sure that these other possible universes are incapable of supporting any sort of life, because the basic structures which can arise in these universe cannot carry physically meaningful information, which is the starting point for life. In other words, the coincidences are sufficiently general that we can be pretty certain that all possible life is ruled out except in the small region of parameter space we happen to reside in, and maybe one or two other similarly improbable pockets where something akin to chemistry can develop in a different way.

So why is this relevant? What if life requires some stupendous chance in the make-up of the fundamental physical parameters to emerge? Can't we just accept this chance, and not worry about it at all? Of course our universe is going to be right for us, since we live here. And what has this to do with metaphysics?

The key is that we have a number of competing models. The first is classical theism, which I will denote as T (for theism), where the physical parameters are ultimately a reflection of God's intentions. The second is the naive mechanism considered around the time of the enlightenment, which I will denote as A, where the physical parameters just happen to be what they are. And there are, of course, other models as well. Now suppose that we specifically consider a form of classical theism where the intentions of God are that life emerges. This seems highly plausible; in the last chapter I argued that God desires that the universe should contain good; and that the existence of rational agents is a good. If we suppose that the existence of material rational agents requires life (which seems certain; a rational agent has an intellect, which is capable of storing information and also communicating it, which are the requirements I used to partially define life a few paragraphs ago), then to create a good universe God would create one containing life, if it were logically possible. In this case, in deism the physical parameters would be 'tuned' by God in His act of creation to within the set of parameters in which life can emerge; in theism we should instead say that the physical constants reflect the immutable will of God, so if God's desire for good requires them to take particular values, then of course they will take those values. On the other hand, in mechanism, the only reason for choosing those parameters over another would be blind chance. They could be anything; they just happen to be right for us. In other words, if T is correct (or deism is correct; but deism comes with numerous problems of its own which theism avoids), then the probability the sort of universe in which we live emerging would be sufficiently close to one that we need not worry about the difference. If model A is correct, then the probability is sufficiently close to zero that we need not worry about the difference. We also denote F as our current theoretical knowledge of physics (excluding those things which are taken solely from experiment, such as the measured values of these fundamental parameters), logic, philosophy and anything else of relevance.

We now unleash Bayes' theorem. We denote $P(L \setminus F)$ as the probability that the the universe is suitable for life to arise (recall that probability is a measure of our uncertainty of a conclusion given knowledge of an insufficient or uncertain set of premises). We know that $P(L \setminus TF)$, the probability that the universe is such that life can occur if theism is correct is close to one. Equally, we know that $P(L \setminus AF)$, the probability that life can occur if mechanism is correct, is exceptionally small. But what we really want to know is the probability $P(T \setminus LF)$, the probability of theism given that the universe is fit for life, in terms of the prior probabilities $P(T \setminus F)$ and $P(A \setminus F)$ based on the various other factors we have considered so far (such as philosophical consistency, and how well the theories are consistent with modern physics). Bayes' theorem gives us the means to calculate $P(T \setminus LF)$

in terms of the other numbers. The equations are,

$$\begin{aligned} P(T \setminus LF)P(L \setminus F) &= P(T \setminus F)P(L \setminus TF) \\ P(A \setminus LF)P(L \setminus F) &= P(A \setminus F)P(L \setminus AF) \\ \frac{P(T \setminus LF)}{P(A \setminus LF)} &= \frac{P(T \setminus F)}{P(A \setminus F)} \frac{P(L \setminus TF)}{P(L \setminus AF)}. \end{aligned} \tag{701.1}$$

Now let us suppose that the prior probabilities of the geometrical and theistic views of physics are the same ($P(T \setminus F) = P(A \setminus F)$). This would be reasonable if they equally explained physics and were equally philosophically powerful. I do not believe this to be the case – in both cases, I believe that theism wins by a considerable margin but for the sake of this argument I will suppose that they have similar prior probability. The probability we should assign to one of the theistic theories being correct compared to the probability that one of the geometrical metaphysical theories being correct is thus close to the ratio $\frac{P(L \setminus TF)}{P(L \setminus AF)}$. We have already stated that $P(L \setminus TF)$ is close to one, while $P(L \setminus AF)$ is exceptionally small. This puts theism as being vastly more likely to the naive geometric view, as long as $P(T \setminus F)$ is not vastly smaller than $P(A \setminus F)$.

Now, obviously this calculation is far too simplistic. Firstly, we do not know what the probability $P(L \setminus AF)$ is, only that it is exceptionally small (i.e. so close to zero that we are perfectly justified in ignoring the difference). We don't even know if it is sensible to apply probabilities to the question of whether theism or atheism are correct (as discussed in section 14.7), although in this case there is more merit in the suggestion, since we are basing the probabilities on a classical interpretation. We divide the region of physical parameters into small volumes. We state that there is a symmetry between these volumes, since all of them are intrinsically possible, and assign the same intrinsic probability to them. We then adjust this probability by either additionally assuming that A is true or T is true, and from this can calculate the probability that the parameters are as we observe them given either A or T and whatever other assumptions we used to divide up the parameter space as we did. Clearly, whatever number we get at the end of it is going to depend on our arbitrary choice of measure. But the fine-tuning argument is that whatever measure we choose initially, as long as it is determined independently of the observed results and not biased towards them, we will find a much higher probability for the values of the parameters we observe if T is correct rather than if A is correct. The difference between this case and that of section 14.7 is that here we are comparing numbers within a mathematical space, while in section 14.7 we were asked to classify theories into sets of theories, a process which can't be so easily quantified.

In principle, the fundamental physical parameters could take any range, up to some very large number. They might be constrained by some self-consistency principle, but if so that principle is still unknown. If we assume that the probability distribution for each parameter is uniform and we assume that the parameters are independent of each other, then the probability would be a small but finite area of phase space over something close to infinity: in other words, near enough to zero that we need not worry about the difference. However, both these assumptions are questionable. We simply do not know either enough about physics or about how it is linked to the metaphysics to even know for certain what the allowed ranges of the parameters are, or what the probability distribution (or the measure) over that range is given the range of metaphysical theories (or even what it means to discuss a probability distribution in this case). Should we describe the possible values of α_W by a uniform probability distribution (which, if α_W ranges from zero to infinity, is mathematically ill-defined), or should we use some function of α_W such as $1/(\alpha_W^2+1)$ or $e^{-\alpha_W}$ (which, if α_W ranges from zero to infinity, are not mathematically ill-defined), or a Maxwell-Boltzmann distribution based on the free energy of the initial state of the universe? The precise number we compute for the probability $P(L \setminus A)$ will depend on this choice, but we haven't a clue where to start choosing the right one. In other words, we are uncertain, which means that

we need to treat the problem of the probability measure itself using probability, perhaps by averaging over all mathematically consistent possibilities for the measure for each of the fundamental parameters (see (Barnes, 2017) for a discussion of this problem).

Now this immediately suggests a question: what if the probability distribution for the values of the fundamental parameters just happens to peak in the right area (which is, essentially, what we are claiming for theism)? This doesn't seem right, though: we would have to ask why it does so, which in itself requires some sort of fine tuning. In other words, on the basis of what we know so far concerning physics, $P(L \setminus AF)$ seems to be small. It might be that as our knowledge of F improves this probability will also change; however right now we can only base our computation on what we know so far. We are right to use F is the prior knowledge rather than some future improved knowledge F', when we don't even know now what F' is. We would have to consider all the possible F's which we both can imagine and haven't yet imagined, and sum them up with weights based on what we know so far; and that 'what we know so far' basically just returns us to F. So, even though it is difficult to assign a number to $P(L \setminus AF)$, we can nonetheless say that it will be much smaller than $P(L \setminus TF)$, even if we can't say how much smaller. But 'much smaller' is all we need for the conclusion to hold. And, whatever measure we use, 'much smaller' will invariably mean smaller by a factor of 10 to the power of some three digit number.

A more serious objection is that I have been treating T and A as though they are the only options. Of course, T and A do not necessarily have to represent single, unique classes of theories: T represents a class of distinct possibilities, as does A; and I have already discussed D or deism as one possible alternative class of possibilities. In principle, we should consider each of the theories that contribute to T and A individually, and weight them against each other with each piece of evidence using the same sort of formulae. So, for example we might have two theories of theism, T_1 and T_2, and while the evidence of fine tuning will improve the credence we give to them compared to some mechanistic theory A_1, one of them might still be significantly worse than the other (and perhaps even worse than A_1) if it is not logically self-consistent or consistent with the form of physical law. In other words, distinguishing different types of theism or the geometrical metaphysics into this model requires no change of methodology, just more work in order to go through all the different cases.

More seriously, we can ask if there is another entirely different option, M? Indeed, there is one class of models consistent with atheism which can accept fine tuning; the multiverse models. This should not be confused with the Everett multiple worlds scenario mentioned earlier. While the Everett scenario might help with some of the fine tuning coincidences, the fundamental physical constants nonetheless remain fixed in it. The multiverse models instead posit that there are a very large (possibly infinite) number of not only possible but also actual universes, with the fundamental physical constants varying in each one. The idea is then that since the vast number of universes effectively sample the phase space of the physical constants, at least one of them will be able to support life, and obviously that would be the corner of the multiverse we find ourselves in. The probability $P(L \setminus MP)$ requires that only one of the infinite number of multiverse universes supports life, and this too would be close to one.

The multiverse seems to me to be the only way one can overcome these odds in an atheistic metaphysics. When confronted with huge odds, one has three and only three general strategies available: to introduce some bias in the sampling technique which targets the desired solution; to rely on an enormously unlikely chance; or to beat the odds for an individual event occurring with an enormously large number of samples. The first option is impossible for the atheist, since if the solution is directed that implies some immaterial will that is external to the material universe, which is precisely what the atheist denies. Thus to say that fine tuning is the result of some sampling bias naturally implies some form of theism or deism. The second strategy, even before we question what we mean by the 'unlikely chance', is only worth considering if no other options are available (which is not the

case, given that option one is live but contrary to atheism), for reasons already discussed. That only leaves option three, which in this case demands the adoption of a multiverse.

There are again various models for the multiverse, some more plausible than others, and each ought to be considered separately (an important question which a multiverse theory must be able to address before it can be considered plausible is to explain why, when a new universe is spawned, the fundamental constants in that universe take some seemingly random value – why aren't they just all take the same or similar value from the theory that underlies the multiverse?). Some arise naturally from some possible unifications of the standard model and gravity, and the anthropic principle has been used to support the idea that these hypotheses are correct. However, not every possible unified theory implies the multiverse – my simple offering above, for example, did not. None of the multiverse models are experimentally testable in any way: all we can observe is our own universe. In other words, the only evidence we can have for the multiverse is if it is demanded by some theoretical reason: all possible expressions for the laws of physics consistent with experiment uncovered for our universe mandate the existence of a wider multiverse. At the moment, only certain forms of string theory might do so, and string theory as a whole is still far from proven, and it is unclear if it can be proved.

The argument from fine tuning cannot be used to support theism ahead of the multiverse, or the multiverse ahead of theism, since they both equally lead to life $P(L \setminus TF) \approx P(L \setminus MF)$. To prefer one over the other requires different arguments, suggesting that $P(T \setminus F)$ is significantly larger or smaller than $P(M \setminus F)$. Since the multiverse and theistic world-views are metaphysical rather than physical theories, these arguments would have to be based on philosophy rather than physics. Most cases I have seen arguing for M have dismissed T based on relatively weak arguments (such as omnipotence conflicting with omniscience; or on straw man representations of T); I answered what I consider to be the least bad of these arguments in earlier chapters, and I know of none which stand up to a close scrutiny. Arguments against M are to my mind stronger; it introduces additional elements of complexity, not only involving the existence of numerous non-observable universes, but also some mechanism which doesn't itself require fine tuning by which the parameters of physics can change from one universe to another. This will require a new set of laws, not describing physics, but this larger space of universes and how one universe can spawn another, and again asks the question of why the laws are obeyed and why those laws and not some other. The multiverse scenario multiplies the questions; classical theism answers the questions we already have. Of course, I should also point out that theism and the multiverse are not exclusive; it need not be a choice between one or the other; one can have both.

15.12. Summary

Combining the multiverse (answering the fine tuning argument) with the Everett multiple worlds scenario (presenting an alternative to an indeterminate final causality; the alternatives are indeterminism or non-locality which creates issues of its own) is the only real hope I am aware of that atheists have to maintain their philosophy given the evidence from physics. Both of these possibilities have various conceptual difficulties and unanswered complications, and neither of these can even in principle be tested experimentally, and nor do they have convincing theoretical arguments demonstrating them without begging the question against God. In other words, to maintain a disbelief in God given our current knowledge, you need to accept various ideas not only without evidence but without even the remotest possibility of finding evidence. The theist, on the other hand, is content with the most obvious and common sense interpretation of modern physics: that we live in a single, indeterminate, universe. That, of course, is not the same as accepting theism, let alone one particular strand of theism. Of course, while the Everett scenario is inconsistent with theism, the multiverse scenario is not, so even if that leg of the atheistic world view survives, it does not rule out theism. I also note the irony: many modern atheists complain

about the supposed lack of evidence for the 'God hypothesis', while in practice to disbelieve in God requires adopting a metaphysics which is (if it is to be consistent with physics) dependent on philosophical positions for which there cannot even in principle be evidence.

The arguments of this chapter do not prove that the God of classical theism is correct; it may well be that a materialistic metaphysics can also imply the laws of physics as we know them while still maintaining the a-causality required to avoid the implications of the classical arguments for God. However, that the laws of physics are probabilistic seems to considerably reduce the possibility of the naive mechanical view of the universe that has dominated philosophy for the past few hundred years, and which was used to supplant classical theism in Western thought and as the motivation for supplanting theism.

If the considerations of the previous chapter are correct, then God has an intellect and will. The remaining point of contention between theists and atheists is whether God is also free to order the universe as He will. If God has this freedom, then theism is correct, and the only question remaining is what type of theism. If he does not, then theism is incorrect.

How can we distinguish between the possibilities? By examining the evidence in an unbiased manner. Let us suppose that both theism and a materialistic metaphysics (to be generous to the atheists) can account for the laws of physics as we know them. There is, however, an important difference: theism allows for special providence, that is that God may act differently on account of some situation in human affairs. In other words, it allows for the possibility for God to have intervened historically. Materialism states that this is impossible, and there can be no break in the laws of physics. We should therefore examine the historical record. If we find something that can only be accounted for by special providence, then theism is correct, and it is likely that this sign will also suggest what model of theism we should accept. If we find nothing, then it does not prove that theism is incorrect (as theism only admits to the possibility of the miraculous, or at most suggests that miracles would occur at some point in human history, but does not demand that any miracle has already happened), but at least allows materialistic explanations to remain plausible. We should, however, neither assume materialism or theism when performing this examination. We must find a neutral way of evaluating the historical evidence. One established miracle is enough to destroy materialism as a valid philosophy.

In practice, of course, it is not going to be as easy as this. The evidence will reduce to eye-witness accounts, either in past generations or our own, and there may be several explanations which may explain these accounts (firstly that they are accurate, but also that they were invented, interpreted incorrectly so that they could be explained by general providence, and so on). Our assessment of the evidence can never be any more than a probability; we must assess the likelihood of the reports circulating had the event occurred as described against the likelihood of those same reports circulating had it not occurred: the likelihood that the witnesses lied or were deceived by an illusion or mental disorder. But that is a subject for the next chapter.

CHAPTER 16

Miracles

16.1. Introduction

DEFINITION 16.1.1. **General providence**:
General providence describes the continual work in upholding the universe by God as though He had no particular interest in any individuals. This leads to the standard description of physics. △

DEFINITION 16.1.2. **Special providence**:
Special providence describes the act of God in upholding the universe when motivated by acting for the good of His rational creation (ourselves, and any other intelligent life forms in the universe). △

The assumption that God has no interest in, for example, human affairs, which was used in the previous chapter, needs to be questioned. Indeed, from the analysis contained within the chapter on ethics, that if anything similar to classical theism is correct, this assumption must at best be only approximate. God is, above all, agape love; He desires for mankind (and any similar creatures elsewhere in the universe) goodness; mankind is not in a state of goodness and cannot obtain it by his own efforts; therefore God has to intervene to provide the best possible future for our race (while still respecting human freedom, since overriding freedom would be a greater evil than continuing to permit freedom to be misused). To derive the laws of physics we assumed that there was no ultimate difference between one place or time and another, but if God is interested in human affairs, then that assumption fails. An unchanging God may act in ways that seem different if the underlying temporal circumstances are different. There might be occasions where the laws of physics seem to be violated.

This is not the disaster that it would be in a mechanistic world-view, where the laws of physics are the ultimate and final step in the ladder of explanations. People often talk about a 'breaking of the laws of physics', but that rather misses the point. Miracles (in the view of the theist) come from the same source as standard physics, only applied to different circumstances, so the effects seem to be different. Rather, it is better to say that the laws of physics are not universal (in the same sense that the mechanist understands them to be) because they are only part of the story of God's interactions with the universe. They describe God's actions in upholding the universe if God has no particular interest in what happens. The full picture, on the other hand, includes both those events to which God is (largely) impartial and those for which He has a distinct purpose to perform, ultimately for our benefit. The laws of physics are thus only an approximation to reality – albeit one which holds in almost all circumstances.

The omnipotence of God should be clear: we started by giving God the ability to destroy the universe at one moment of time, and recreate it in any form; saving ourselves from

anarchy only by insisting that that this should be according to some probability amplitude constrained by what we know of the nature of God including His rational nature. The derivation of that probability amplitude, however, breaks down in those specific cases where God has a particular purpose. We can then expect that something out of the ordinary will happen. If it is required for God's overriding purpose (which is, of course, agape), there is nothing to stop God from reversing entropy, increasing the surface tension of water to allow a man to walk across it without sinking, initiating a conception within a virgin, or raising a man from the dead; and the way God does so is no different from the way in which He guides matter in the ordinary course of things. It only seems different to us because it seems to violate the usual regularities, but those regularities are drawn from an incomplete picture of the universe; the complete picture would include God's special providence, in which case miracles would be regular and understandable. However, since we cannot understand the will of God nor subject it to experimental test, they lie outside the purview of science. That doesn't mean that they are wrong (not all truth is scientific truth); just that different tools are needed to understand them. A miracle is thus not a breaking of the 'laws underlying the laws of physics', but instead both fully consistent with them and outside their purview, and could be fully expected and understood if we held a complete understanding of the mind of God.

The difficulty is that we don't know what correction is needed to the mathematical detail of fundamental physics to account for special providence; we do not know the mind of God well enough, and even if we did, the correction to the premises would be related to the natural purposes of complex living organisms, and getting from that down to the level of fundamental physics is very challenging. Not impossible: it can, of course, be done, just not by us at the moment with our limited brain power and computing resources. A full first principles mathematical treatment of miracles is thus possible in principle but exceptionally difficult (and possibly impossible) in practice; and that restricts us to a somewhat ad-hoc treatment which makes it seem as though miracles are a violation of the natural regularities of physics and metaphysics. It is not the miracles themselves which are a violation, but only the crude way we are forced to treat them makes it seem so because we only know the effects and not the complete fundamental theory to adjust our premises. We cannot merge our account of miracles into physics even on a full account (and claim that they thus natural events and consequently can be accounted for without reference to God), because physics is itself dependent on God, and physics describes the regular theory based on the approximation that God is impartial to mankind. Once we relax that approximation, we move beyond physics to incorporate aspects of theology as well, but the combined approach would make the overall analysis more complete and thus more accurate, if we could only do it.

DEFINITION 16.1.3. **Miracle:**
A miracle is an individual act of special providence, namely an event, which, if it truly occurred, when taking into account what occurs in apparently similar circumstances, provides evidence for the contention that the basis underlying the physical world is not wholly impartial to the concerns and needs of rational life forms. Δ

Discussion:
The phrase 'if it truly occurred' is intended to exclude any complications caused by the evaluation of the evidence suggesting whether or not the purported miracle actually happened. Such discussions are important, but to avoid any circularity when we discuss the observed evidence for miracles, such considerations need to be excluded from the definition of a miracle. The phrase 'taking into account what occurs in apparently similar circumstances' is to avoid those events which by coincidence seem beneficial. For example, if we had an event with a one in a thousand probability, observation of this event without consideration of

what happens elsewhere might seem to provide evidence for God's actions on our behalf. However, if we also include the nine hundred and ninety nine cases where the event did not occur, then in total there is no evidence that the universe is not impartial to us; indeed the totality of evidence in this case would suggest that it is. We must therefore consider the totality of evidence when identifying something as a miracle. On the other hand, those things which are not only improbable but impossible in general providence, such as those that violate the conservation of momentum or the second law of thermodynamics, can be clearly identified as miraculous (if they occurred), since they would demonstrate that there is more to God's upholding the universe than just general providence[1]. If what ought to be a one in a thousand chance (according to accurately reasoned predictions based on firmly established theory built on the assumption that physics is impartial to us) in fact turns out to occur one time in a hundred, then the excess events might provide evidence for some form of theism, if the statistical sample is large enough and the calculations were correct and based on the correct theory. By the basis underlying physics I mean whatever it is that explains why physical law takes the form that it does, whether it that physical law is a description of God's upholding of the universe or some other explanation (including that there is no explanation, an idea which would count as one of the bases which are wholly impartial to us). I say 'wholly' impartial because that is what those anti-theistic metaphysical theories miracles are purported to provide evidence against demand. There is nothing in theism that is contrary to the idea that the basis underlying physics seems to be largely impartial to our needs and concerns; only that we should expect to see a non-zero amount of direct evidence of God's concern for us. Fine tuning arguments do not count as miracles under this definition as they concern the form of physical law rather than specific events.

I should observe that this is not the Biblical definition of a miracle. In the Bible, several words are used to denote miracles: one is that they are a sign, secondly they are referred to as works, thirdly as something which induces a sense of wonder. Clearly, an individual act of special providence invokes wonder, and a sign providing evidence for an aspect of God's character will also always demonstrate His lack of indifference. Thus my definition is equivalent to the biblical one, in that the signs of the Bible would be classed as miracles by my definition, and all my miracles would be called as signs in biblical language. However, my definition does miss something of the Biblical description. Firstly, signs almost by definition point towards something. They are not arbitrary, but have a clear purpose. In particular, the Biblical signs go beyond merely showing God's indifference, but exist to deepen the relationship between God and man. Secondly, we see that miracles are not only a matter for intellectual inquiry, but designed to produce an emotional response. They bring us closer to God in both our minds and our passions, which is, of course, as God intends.

I should also say that my definition is limited in that it does not include acts by supernatural but non-divine agents (if any exist). This keeps things simpler, but is limited; we may well treat (for example) an act of a demon as miraculous, or it may be indistinguishable

[1]Note that the conservation of momentum emerges from translation symmetry in classical mechanics, and locality in quantum field theory, both of which arise from the similar concept that the interactions affecting a particle only depend on the behaviour of the fields at that location. My derivation of the tendency towards equilibrium also assumed locality, i.e. the evolution of the physical states only depended on the state of the universe at that time (and space, once we include special relativity). I explained the concept of locality from the idea that for God there is no preferred location or time in the universe, so only relative displacements could affect God's interaction with the universe, and any relative displacement of a continuous function can be reduced to an expression built from point-like differential operators. This is directly based on the assumption that God is impartial to the various events happening in the universe. Locality is, however, violated in a miraculous situation if God is responding to prayer or preparing for some future event. In this case, an event at an absolute location and time (whether the prayer or the future event) can influence God's interaction with the world. Thus violations of the conservation of momentum and the second law are only in contradiction to the premise that God is impartial to our part of the universe, which we are only treating as approximate in this chapter, and not to God's nature or the idea that God sustains the universe.

in its effects from an act of God. Of course, it is important that we do distinguish them; but this is a difficult question, and better addressed in a more complete discussion of miracles. □

Now the next question is why (if theism is correct) miracles seem to be so rare. In some distant corner of the universe, the answer is straight forward: matter progresses as though God had no special interest in its destination because God does have no special interest in its destination. God's primary purpose in creation was to pour out agape on rational agents such as ourselves, so if something doesn't affect us, then there is no reason for any special providence. This begs the question of why create such a vast universe, if all God needed was the Sun, Moon and Earth? One answer to this question is simply to ask in response 'Why not?' The presence of other planets, stars, and galaxies makes no difference to God's relationship to ourselves (and any other life forms elsewhere in the universe). If, on the other hand, God's preferred mode of operation was to use a big bang, inflation, star formation, evolution by natural selection, and so on, all acting according to General providence, to obtain His goal, then what disadvantage does this have to God over an immediate special creation? Indeed, if 'The heavens declare the glory of God, and the sky above proclaims his handiwork[2],' then this could hardly be the case if there were no heavens to do so. The stars have inspired mankind in numerous ways, for poets no less than modern physicists. To state that God would have been better off creating just 'what He needed' presupposes that we ourselves would have benefited more from the simplest possible specially created universe rather than one that has the appearance[3] of one which developed from General providence; an unproven and possibly unprovable presupposition[4]. Why would God wait through all the billions of years until mankind emerged? This objection, a form of one of the oldest objections to Christianity[5], ignores the timelessness of God. God does not wait; to Him a billion years are the same as a day, and a day the same as a billion years.

But why then do miracles seem to be so rare in our particular corner of the universe? Here, I think that the answer is two-fold, both relating to human freedom. The first is that freedom is only possible against a background of regularity. If God were constantly interfering, that would include constantly interfering with us. That would prohibit human freedom, which includes freedom to act contrary to God's specific will[6]. This requires a certain independence from God, which is best achieved by God acting as though He were impartial to our choices, leading to the regularities which lead to the laws of physics. Secondly, in the particularly Christian account, at the beginning of human history our ancestors were given a choice: either to live in a world where God would continually act specially to preserve their good, but on the condition that they were nice to each other; or to have the freedom to be nasty to each other and seemingly live independently of God. In the spirit of all teenagers since then, our ancestors replied (in words to this effect) 'Up yours, Daddy, we don't need you.' And God replied, 'OK. Things will be tough for you without my protection: you will be under the curse that you will face thorns and weeds

[2]Psalm 19:1.

[3]To be pedantic, the most we can say is that the universe has the appearance of having developed over a long physical process. However, on the basis that God is not a liar, it is very likely that this appearance also reflects reality.

[4]Of course, my claim that the universe we see is better for us in the eyes of God than some minimal universe is equally unprovable; but that misses the point: while it is possible that God would have preferred to have created the universe as it appears to be then the objection that he would be more likely to have created something simpler loses its force.

[5]Irenaeus, *Against the Heretics*, 2.28.3; Augustine, *Confessions*,11.10-14.

[6]As with general and special providence, it is usual to distinguish between different desires of God. For example, God desires that all should repent and be saved from perishing, but equally He desires to respect our freedom of choice. There is then within God two senses of desire: this is how I would prefer for the world to turn out, but, given my respect for human choices, this is the best that I can in practice do. We might think of the difference between long term hopes and aspirations, and our day-to-day pragmatic choices which need not always coincide with our dreams and hopes.

and have to sweat for your bread; but I'll stand back until you or your descendants change your mind or your nature becomes so corrupted that you are on the point of complete self-destruction.' And that same response to God has been repeated in every generation since (not, though, through our choice: we are conceived into that state of rebellion), with only a few notable exceptions who have rejected that rebellious attitude and submitted to God (most importantly Abraham and Mary[7]) through whom God has been able to prepare for his eventual purpose of salvation. This is, of course, a somewhat stylised and over-simplified reading of the first few chapters of Genesis; there is far more to the story than just that; but, while the dialogue above is my own personal embellishment, the idea that we are by the choice of our ancestors and through pride rebels against God and that is the reason we see all this suffering around us is central to all forms of orthodox Christianity. Now note that, in particular, standing back would have the appearance of indifference to the fate of humanity: otherwise God would either be acting for us (to provide miraculous benefits) or against us (to provide miraculous punishment). And this seeming indifference leads, as we have seen, to a world governed by general providence and physics. Not that God is truly indifferent; He (speaking metaphorically) weeps for every one of us who desires to continue in rebellion to our own destruction; but nonetheless, to frequently act otherwise would be to destroy our freedom.

Of course, I have not proved that such an account is true; but my intent was description rather than proof. If Christianity (or some similar religion) is true, then we expect to see a world which is mostly regular in its operation, baring a few rare exceptions in response to some expression of faith which are consistent with God's agape for both individual people and mankind as a whole, and consistent with some overall plan for our rescue from sin. Miracles should not be some arbitrary and whimsical displays of power (as we see, for example, in the gnostic gospels, Islamic hadith, and fairy stories), but each should be part of a larger story of God's agape for mankind.

Note also that those who believe in miracles must have a strong belief in the otherwise regular nature of the universe. One cannot have exceptions to a general rule if there is no general rule. One cannot identify something as an unusual act of God if that is the sort of thing that happens usually. If you want to undermine Christianity through physics, you are far better off trying to show that the resurrection of Christ could occur through the normal course of events than showing that it could not. That would leave the historical evidence intact, whilst undermining the event's meaning and significance.

In a mechanical view of the universe, of course, such exceptions are impossible. In this philosophy, the universe appears impartial to Mankind because that which provides the foundation for the laws of physics is wholly impartial to mankind.

16.2. Objections to the Miraculous

Objection:
Many people object to miracles on the basis that they are outside our normal realm of experience; the claims are apparently extraordinary and so the evidence needed to support them ought to be equally extraordinary. One of the most well known and well presented arguments of this type was by David Hume[8].

> A wise man, therefore, proportions his belief to the evidence. In such conclusions as are founded on an infallible experience, he expects the event with the last degree of assurance, and regards his past experience as a full proof of the future existence of that event. In other cases, he

[7]Genesis 15:6; Luke 1:38.

[8]I am here quoting the argument as it appears in Hume's *Enquiry concerning human understanding*, Book X.

proceeds with more caution: He weighs the opposite experiments: He considers which side is supported by the greater number of experiments: to that side he inclines, with doubt and hesitation; and when at last he fixes his judgement, the evidence exceeds not what we properly call probability. All probability, then, supposes an opposition of experiments and observations, where the one side is found to overbalance the other, and to produce a degree of evidence, proportioned to the superiority. A hundred instances or experiments on one side, and fifty on another, afford a doubtful expectation of any event; though a hundred uniform experiments, with only one that is contradictory, reasonably beget a pretty strong degree of assurance. In all cases, we must balance the opposite experiments, where they are opposite, and deduct the smaller number from the greater, in order to know the exact force of the superior evidence.

To apply these principles to a particular instance; we may observe, that there is no species of reasoning more common, more useful, and even necessary to human life, than that which is derived from the testimony of men, and the reports of eye-witnesses and spectators. This species of reasoning, perhaps, one may deny to be founded on the relation of cause and effect. I shall not dispute about a word. It will be sufficient to observe that our assurance in any argument of this kind is derived from no other principle than our observation of the veracity of human testimony, and of the usual conformity of facts to the reports of witnesses. It being a general maxim, that no objects have any discoverable connexion together, and that all the inferences, which we can draw from one to another, are founded merely on our experience of their constant and regular conjunction; it is evident, that we ought not to make an exception to this maxim in favour of human testimony, whose connexion with any event seems, in itself, as little necessary as any other. ...

But in order to increase the probability against the testimony of witnesses, let us suppose, that the fact, which they affirm, instead of being only marvellous, is really miraculous; and suppose also, that the testimony considered apart and in itself, amounts to an entire proof; in that case, there is proof against proof, of which the strongest must prevail, but still with a diminution of its force, in proportion to that of its antagonist.

A miracle is a violation of the laws of nature; and as a firm and unalterable experience has established these laws, the proof against a miracle, from the very nature of the fact, is as entire as any argument from experience can possibly be imagined. Why is it more than probable, that all men must die; that lead cannot, of itself, remain suspended in the air; that fire consumes wood, and is extinguished by water; unless it be, that these events are found agreeable to the laws of nature, and there is required a violation of these laws, or in other words, a miracle to prevent them? Nothing is esteemed a miracle, if it ever happen in the common course of nature. It is no miracle that a man, seemingly in good health, should die on a sudden: because such a kind of death, though more unusual than any other, has yet been frequently observed to happen. But it is a miracle, that a dead man should come to life; because that has never been observed in any age or country. There must, therefore, be a uniform experience against every miraculous event, otherwise the event would not merit that appellation. And as a uniform experience amounts to a proof, there is here a direct and full proof, from the nature

> of the fact, against the existence of any miracle; nor can such a proof be destroyed, or the miracle rendered credible, but by an opposite proof, which is superior.
>
> The plain consequence is (and it is a general maxim worthy of our attention), 'That no testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous, than the fact, which it endeavours to establish; and even in that case there is a mutual destruction of arguments, and the superior only gives us an assurance suitable to that degree of force, which remains, after deducting the inferior.' When anyone tells me, that he saw a dead man restored to life, I immediately consider with myself, whether it be more probable, that this person should either deceive or be deceived, or that the fact, which he relates, should really have happened. I weigh the one miracle against the other; and according to the superiority, which I discover, I pronounce my decision, and always reject the greater miracle. If the falsehood of his testimony would be more miraculous, than the event which he relates; then, and not till then, can he pretend to command my belief or opinion.
>
> ... Upon the whole, then, it appears, that no testimony for any kind of miracle has ever amounted to a probability, much less to a proof; and that, even supposing it amounted to a proof, it would be opposed by another proof; derived from the very nature of the fact, which it would endeavour to establish. It is experience only, which gives authority to human testimony; and it is the same experience, which assures us of the laws of nature. When, therefore, these two kinds of experience are contrary, we have nothing to do but substract the one from the other, and embrace an opinion, either on one side or the other, with that assurance which arises from the remainder. But according to the principle here explained, this substraction, with regard to all popular religions, amounts to an entire annihilation; and therefore we may establish it as a maxim, that no human testimony can have such force as to prove a miracle, and make it a just foundation for any such system of religion.

The thrust of the argument is that we must judge a historical event using a probabilistic method: evidence counted for it, such as eyewitness testimony, can in principle raise the prior probability of an event to the sort of level where we are justified in believing that that event occurred. However, no eye-witness testimony is infallible: people lie, or they might be mistaken, or their testimonies distorted through transmission across several generations. The final probability we accord to an eyewitness account depends as much on the probability we assign to such an event before we consider the direct evidence. If that initial, or prior, probability is exceptionally low, it must takes a very secure degree of testimony to make us believe it. If the supposed event is deemed impossible, then no amount of testimony (as long as that testimony is less than certain) would be enough to persuade us that the event occurred.

So how do we calculate the prior probability? According to Hume, probabilities are measured by repeated observation. So, if we observe a set of circumstances N times, and see that outcome X occurs n times, we can say that the probability of X occurring is close to n/N, with the precision of this estimate being better the more frequently we observe the event. So, for example, if we uniformly observe in observations we take to wholly reliable that when people try to walk across water they sink, and we observe this a large number of times, then we can say that the probability of someone successfully walking across water is, if not zero (since this sort of argument can't provide the infinite precision needed to state this), at least so vanishingly small that we can for all practical purposes forget about it.

This means that when we come to some ancient testimony that somebody did walk across water, the chances of the witness being incorrect, no matter how small they seem, are still immensely greater than the prior probability we have calculated for such an event occurring. Therefore their testimony should be disbelieved.

A miracle is defined (by Hume) as a breaking of the law of nature. Through an almost immeasurable number of observations we can say that the probability of the laws of nature being broken is exceptionally small. No human testimony or experience, even our own, is enough to overcome this probability. Therefore every account of a miracle ought to be disbelieved. •

Response:
Hume's argument is one of those which seems convincing until we stop to think through the details. Hume makes a number of unstated assumptions and methodological mistakes which, on reflection, cannot bear the burden he is placing on them. This was pointed out even before Hume wrote (he adapted his argument from earlier deist writers, who had already been refuted by the time Hume put pen to paper, though Hume gave no credit to his predecessors nor mentioned the objections of their detractors).

Note that Hume makes an error in the first paragraph I quoted, saying that we side with the theory that is supported by the *greater number* of experiments. On the contrary, the true theory should be supported by *all* experiments. Of course, there is more to it than this: we need to take into account systematic and statistical experimental error, and discard those experiments whose set-ups were sufficiently flawed that their results are meaningless. If the theory makes probabilistic predictions then we need to compare the distribution of the totality of experimental results against the theory, which complicates matters. But what Hume is suggesting is different – he is trying to choose between two deterministic theories. In fact, it is the exceptions which are usually most interesting. But every experiment which is discarded must have a valid reason for letting it go; that it is part of the minority is not a valid reason. Hume thus starts out by showing that he doesn't understand the experimental procedure. He makes no mention of experimental errors nor shows that he appreciates their importance. In practice, the errors often make the results of experiment ambiguous, with numerous possible theories in agreement with them.

Secondly, Hume assumes from the start the regularity of the universe. *In such conclusions as are founded on an infallible experience, he expects the event with the last degree of assurance, and regards his past experience as a full proof of the future existence of that event.* But this is only true if the universe is known in advance to be perfectly regular, and in particular if the type of regularity is as supposed by Hume (note that the interpretation of what past experience has to say about future existence is model-dependent, and thus nowhere near as simple as Hume supposes). But that is just what we are investigating in the case of miracles. Thus Hume's argument is circular.

It will be sufficient to observe that our assurance in any argument of this kind is derived from no other principle than our observation of the veracity of human testimony, and of the usual conformity of facts to the reports of witnesses. Once again, Hume rests his argument on an at best over simplified and at worst false premise. We know that human testimony is sometimes reliable and sometimes not; we know the circumstances under which people lie, and under which they tell the truth, and under which they might mis-remember or be deceived and get some details wrong. And how do we know this? Because we ourselves are human beings, and have given testimony on many occasions. Thus by omitting the various factors that affect how we judge human testimony, Hume dangerously oversimplifies his case. Furthermore, observation is not our only means of testing human testimony. Additionally, for many of the things people testify about, that testimony is the only evidence that we have for the truth of falsity of the events. So how are we to ascertain whether or not the testimonies are true, if they are the only evidence available? Hume is therefore wrong

in saying that our only guide is observation, wrong in not considering how circumstances weight the reliability of testimony, and wrong in his understanding of how we understand truth. Experience as Hume understands it is not the only means we have of establishing truth, and because his argument relies on his epistemology, Hume's flawed epistemology undermines his whole project.

A miracle is a violation of the laws of nature. But Hume's definition is only useful if it is one that theists would accept. That atheists accept it is of no consequence: they are not the ones who need convincing. But it is clear that few theists would accept this definition. Hume interprets *laws of nature* from the perspective of an empiricist. He takes a bunch of experimental data, draws a line through it, and that's his law of nature. The argument also works with the idea of mechanical laws of nature. But theists accept neither mechanism nor empiricism. Since theists don't believe in laws of nature in the same way that mechanists or empiricists do, they can't define a miracle as a violation or breaking of them. Since mechanism wasn't developed until the fourteenth century, and empiricism later than that, it is clear that no Christian in the first thirteen hundred years of the religion, or Muslim in the first seven hundred years of theirs, could have accepted this definition of a miracle. Instead, a theist would define a miracle in a similar way to how I did above. This is a major problem for Hume. Even if his argument is successful, he hasn't disproved the miraculous as the most important authorities in either religion would have understood the term. There might be events which are miraculous according to the theist definition, but not according to Hume's; and Hume would not have disproved that such events have occurred.

But now we come to the centre of Hume's argument, and here the unease about it only increases.

Perhaps the first sense of unease comes from when we consider what Hume is asking us to do. We are, essentially, being asked to judge observational evidence on the basis of a theory. That that theory is ultimately derived from induction and thus observation itself does not change the overall picture. Essentially the procedure that Hume is advocating is that we make a number of observations; construct a model of how the universe behaves on the basis of those observations; and then judge whether further observations should be believed on the basis of that model. That Hume did not discuss the modelling step of his argument does not mean that it is not present. Indeed it has to be present: we can say that experiment contradicts theory, because a theory makes a general prediction across all space and time including when we conducted the experiment; however, one event (and an experiment is the record of an event) at one place and time cannot contradict an event at another place and time. One cannot say that because a dog barked on Monday that he can't growl on Tuesday. One can only say that the dog can't bark and growl at the same instant of time. If one day I perform an experiment and it responds that ten particles were produced, and the next day I perform the same experiment and fourteen particles are produced, there is no contradiction. The only contradiction is with a theory that states that I must get the same result each time I perform the experiment. In other words, experimental evidence alone cannot make a prediction – only theory can predict. But to have a contradiction requires that we have a disagreement between an observation and a prediction. So when Hume is making his measurements, noting down regularities, and using these regularities to derive statements of probability, he is constructing a model or theory. And make no mistake – 'Things happened a hundred years ago in much the same way that they happened today; this is what happens today, therefore this we expect to have happened a hundred years ago,' is not simply an experimental report but a model built on the experiment and an additional assumption.

So what Hume is in effect asking us to do is to assess the validity of an observation on the basis of a theory – but this is the very opposite of what scientists are trained to do: we judge the validity of theories on the basis of experiment. Of course, it is worse than this: Hume's model is not merely constructed on the basis of observation, but observation plus

some unstated assumptions, most notably that external circumstances are the same when the supposed miracle occurred as when we performed the experiments that we use (in part) to build our model. But it is precisely these assumptions that the account of the miracle is put forward to challenge. Again we see that Hume's argument is ultimately circular: when constructing the model it makes certain assumptions, and then dismisses evidence that might be used to question those assumptions on the basis that they contradict the model.

A second sense of unease comes from Hume's treatment of probability. We have to be a little bit careful here in the terminology; we cannot say 'the probability of an event having occurred' because past events in themselves can't carry any probability other than 1 or 0: either they happened or they did not. Rather, we should discuss how certain we can be given the available evidence and the possibly numerous different models which could be used to interpret that evidence. For example, if we consider a certain event M, and whether it is likely to have occurred in the face of some new evidence E and old evidence W. We are therefore interested in the probability of M given E and W, which I will write as $P(M \setminus E, W)$. From a simple application of Bayes' theorem, we find

$$P(M \setminus E, W) = \frac{P(E \setminus M, W)}{P(E \setminus W)} P(M \setminus W). \tag{714.1}$$

This basically says that the chances of M having occurred given the evidence in E and W is proportional to the chances of M having occurred given the evidence of W times some quotient representing the probability of the evidence having occurred both if the miracle occurred and in general (whether or not the miracle occurred). We can also easily derive from this equation the ratio of the probability we should assign to our belief that the miracle occurred given the evidence to that it did not occur given the evidence, and this is basically the number we are most interested in.

$$\frac{P(M \setminus E, W)}{P(!M \setminus E, W)} = \frac{P(E \setminus M, W)}{P(E \setminus !M, W)} \frac{P(M \setminus W)}{P(!M \setminus W)}. \tag{714.2}$$

This removes the quantity $P(E \setminus W)$, i.e. the probability of the eyewitness testimony given our background information, which is difficult to assess. If the right hand side of this equation is infinite, then we can say with complete confidence that M occurred; if it is zero then we can say with confidence that M did not occur.

The right hand side of this equation contains two ratios. The final term is the ratio of the probability that M occurred given our knowledge of W to it not occurring; the first term gives the effect that the eyewitness testimony has to change this number: it gives the probability that the eyewitness account was given if the miracle occurred to the probability that such an account would be given if the event did not occur. This is where our knowledge of the eyewitnesses comes into play; for example if we have a habitual liar who would receive a great deal of personal gain if his account was believed, then the probability that he would present the testimony even if the event did not happen would be quite high. If, on the other hand, we had a rigidly honest person whose interests suffer from the account, then the chances of him making up the testimony if the event he claimed to have witnessed did not occur would be very small, and this evidence ratio would be very large; we should believe that M occurred even if it seems to have a low intrinsic probability in the absence of the evidence. In a modern rephrasing of the argument, we can say that Hume's assertion is that in the case of miracles, the ratio $\frac{P(M \setminus W)}{P(!M \setminus W)}$ is so small that no matter how unlikely it is that the witnesses would have just made it up it can't bring the overall result up to some value where we might consider believing the story. Of course, in any historical investigation, it is difficult to assign numbers to these probabilities (once again, somebody having lied is an event and we can't assign probabilities to past events); but nonetheless this formula is useful in getting a rough order of magnitude estimate.

Of course, Hume was ignorant of this mathematical structure: it was developed well after his time. And the argument as quoted above displays this ignorance. Hume tries to argue that we generally believe eyewitness testimony on the basis that it has repeatedly been shown to be reliable; this is, however, false. We generally believe eye-witness testimony in part because the probability that that testimony appearing if the event did not occur is usually much less than if it did occur, unless there is a good reason why the witness might want to lie. People don't generally make up strange stories they were a witness to and then try to pass them off as fact. For example, suppose we found in a historical record a statement made by an indifferent witness 'The Queen addressed the troops wearing a white dress.' Why would he say a white dress? Perhaps the witness just made up this detail, but in that case why say white rather than red, blue, or black? If he was just making it up, then he could have picked any colour. If he did not know or could not fully remember then he could just omit the detail altogether, and it would likely make little difference to his overall account. On the other hand, if he was there and the dress was indeed white and its colour particularly stuck in his memory, then he might well mention the detail. The probability of him saying a 'white dress' is therefore vastly greater if the Queen indeed wore white (or some colour that could easily be misidentified as white) than if the Queen did not. Of course, the historian always has to take into account the possibility of bias or poetic embellishment, but in general if somebody made up some story and passed it off as fact there is always the issue of 'Why that story, and not another?' which makes the chances of it being true vastly greater than it being false under normal circumstances.

Hume's view of probability is, of course, frequentest, and this is his main problem. Frequentest probability is useless when it comes to a discussion of assessing the chances of unique events. Indeed, when assessing our knowledge of events (which is what we are doing), we deny a frequentest interpretation of probability, because we are necessarily presupposing a Bayesian interpretation.

For example, suppose that we were to discuss whether Julius Caesar was assassinated before the Senate of the Republic on the Ides of March 44 BC. We might try to assess this probability in a frequentest way. The Roman Republic lasted for some 482 years, around 176000 days. Let's take the first 88000 days as a sample: that should be large enough for any reasonable assessment. First day of the republic, was a Roman dictator assassinated before the senate? No. Second day, was one assassinated? No. And so on, until we reach the end of the sample, and find that never was a senator assassinated in that sample. We therefore conclude that the probability of an assassination occurring on a specific day, $\frac{P(M\backslash W)}{P(!M\backslash W)}$ is so small that we need not bother about it. Of course, we would have to discuss how reliable the historical sources are - or do we? Hume did not feel that he needed to, so why should we? And even if we do, we all know that history was written by the victors; surely there was motivation for Octavian and Anthony to invent stories of this event (if it did not occur) after the fact to justify their rebellion against the senate; the witnesses cannot be wholly reliable; surely not reliable enough to overcome those odds. Therefore, no matter how good the evidence is that Caesar was assassinated on the Ides of March (rather than, say, one day previously), we should dismiss it as being not good enough to overcome the initial probability; if the frequentest account is correct. That this argument comes to the wrong conclusion gives us good reason to suppose that the frequentest account, as Hume uses it, is not correct.

Critics might interject that political assassination is not against the laws of nature, so this is not a fair example, but that misses the point. Hume claims that we can assess the probability from merely repeated observations of similar events. It doesn't matter whether the object of our study is supernatural or natural. Did I specifically choose the sample to exclude the actual assassination? Certainly, but again, how is this different from Hume; who excludes the account of the miracle from the data he gathers to assess its inherent probability? When Hume states 'But it is a miracle, that a dead man should come to life;

because that has never been observed in any age or country,' is he not excluding from his sample precisely those testimonies under question and which deny his thesis? Is he not assuming what he set out to prove?

Is the data sample I am using sufficiently large? Surely the sample of events which Hume relies on is much larger than the number of days of the Roman republic. 88000 days is a lot, but surely the evidence against miracles is much greater than this? So maybe 88000 days is insufficient to balance the eyewitness evidence, but the experience of non-miraculous events is sufficient to overcome the eyewitness evidence for the miracle? If you like, you can extend the sample to every day in human history except the ides of march 44BC, or every hour except the hour when Caesar was assassinated; and as long as you exclude the actual event (and any similar ones) from your sample (as Hume deliberately excluded the evidence for the miraculous from his sample) you will find that no dictator was assassinated before the Roman senate on those days. The frequentest account simply cannot be used to assign probabilities for unique historical events in the way that Hume wants it to. The point is that no matter how much evidence we have that there were no miracles at other times, that doesn't rule out the possibility that something different happened at the time when the miracle occurred. If the miracle did occur, then you would still have the same other non-miraculous events at other times. We can't draw a meaningful probability about whether or not one particular event occurred based on what happened at other times without making further assumptions connecting those events — assumptions which a theist would dispute. Brutus didn't bring a knife to the senate on the ides of March 45BC. What evidence does that give us about whether or not he did so on the ides of March 44BC? None whatsoever. To turn it into evidence, we need to assume some model connecting the two days. Maybe Brutus never brought a knife to the forum. Maybe he did so only when he planned a political assassination. Both models are consistent with the evidence of 45BC. So that evidence might rule out some models of Brutus' behaviour, but it says nothing about his behaviour in 44BC. The same is true about Brutus' behaviour on any other day (excluding the ides of March 44BC); none provide the evidence we need to judge the probability of whether or not he brought a knife on the day of Caesar's assassination. If there are N days when Brutus didn't bring a knife, and each one provides zero evidence about whether he came with a weapon on the fateful day of Caesar's assassination, then the sum of zero evidence on each of those days totals to zero evidence. Even if we take the limit $N \to \infty$. An non-miraculous event by itself cannot provide any evidence as to whether a miracle occurred at a different place and time. You can't use it, or any number of such events, to construct a probability. To connect the various events, you need a theoretical model. You can then judge how probable the model is on the basis of your events, and then decide how probable the event in question is by its consistency with the model. But the probability for the miracle will be conditional on the model you choose. It is not objective.

Or suppose that we have an event that in practice occurred exactly once in an exceptionally large number, N, of samples (although we don't know that it must occur once, but only that it is rare and might occur). The probability that it occurs in any one sample is thus $1/N$; perhaps this was assessed by some careful assessment of some parallel set of data concerning which we have full knowledge. We now have a testimony, which we judge to be 99% reliable, that it occurred in sample 215715723176; and this is the this only record that we have. Applying Hume's logic, however, we would have to say that the probability that the event occurred in that sample is scarcely any better than any other sample; we have to include the prior probability of $1/N$ and that though the doubt concerning the testimony is only 1%, compared to the prior probability this testimony is nowhere near good enough to make us view that the event occurred. Therefore, we conclude that the event likely did not occur in sample 215715723176. And, of course, we have no good evidence to suppose that it occurred in any other sample either. So we conclude that it did not happen at all. Which is the wrong answer. The reason, of course, that we got the wrong answer is that

the methodology is wrong. We cannot use frequentest statistics to assess the probabilities of specific events, especially rare events, occurring in the sort of way that Hume demands of us.

Now I must come back to the importance of models. What is W in equation (712.2)? I have been writing as though it were a body of evidence, which is, of course, what Hume himself put in there. But this is bad; as I have explained, we cannot compare experiment with another experiment and make a statement of probability; only a comparison between experiment and some model (which might be partially justified by comparison with experimental, or induced from experiment) is allowed. If W represents the body of observed evidence, then we should instead write $P(M \setminus W_1)$ and so on, where W_1 represents some model we have constructed which explains W. Of course, W_1 is not just constructed on the basis of the physics, but also on the basis of various metaphysical presuppositions.

For example, a simplistic interpretation of agape demands that we give food to the starving man and remove the excess food from the glutton[9]. Now, suppose that we come to a town with twenty thousand and one inhabitants. Twenty thousand of these people are starving. The remaining inhabitant is a glutton and has excess food. Motivated by agape, we give food to the twenty thousand people, but when we come to the glutton we do the opposite, and take away his excess. Now the Humean, of course, looks at the sample of a small number of these events to try to predict the probability that we either give or take away food from someone he doesn't yet know about. He first of all looks at our treatment of (say) the first thousand starving people, and constructs a model based on these observations from which he can try to make future predictions (including when we come to the glutton). Now, there are at least two models which explain our behaviour (there may, of course, be others as well, but for simplicity let us stick to these two): firstly, we may be giving food to everyone regardless of their status; secondly we are acting in accordance with agape and ensure that the amount of food that everyone has is within the range of what is healthy for them. Both of these are in complete agreement with the evidence observed so far. But when it comes to the glutton, the prediction is very different: one model states that there is a very high probability that we give him food, and the other that there is a very high probability that we take it away. The first model is in many ways simpler: it assumes that all men are basically equal, and therefore our treatment of them should be the same. But it is wrong. Hopefully the analogy is clear: if we only have induction available (and this is what Hume's argument relies on), then several different models might be consistent with the data we use to construct the models, but the predictions made by the models might be very different. We have no way to test between these models based on the original data: we ought to admit them both as possibilities.

I see the person who assumes that we are giving food to everyone and so by induction concludes that we will also give it to the glutton as equivalent to the Humean who views miracles as for all practical purposes impossible on the basis of his observation of various non-miraculous events. He insists that his own model is the only possible one; but that insistence is begging the question. To make his prediction, he assumes a materialistic, possibly even mechanistic, picture of the workings of the universe, and dismisses the theist picture despite that it fits the original data just as well, or possibly better. The theist on the other hand, looks at the miracle and sees it as entirely consistent with the non-miraculous events: both being examples of how God displays his unchanging nature of justice and agape. We cannot judge between these models (if they are both internally self-consistent and so on) except by looking specifically at those cases where they make different predictions; after this, we will know, if the precision of our observations when the models contradict each other is good

[9]I say simplistic because, though agape desires that the glutton gives up his gluttony, I am far from convinced that stealing from him is the most compassionate way to achieve that end. Nonetheless, for the purposes of this illustration, I will neglect this quibble.

enough, that either the model that contradicts the data is wrong (or perhaps that they are both wrong). But this is precisely what the Humean denies us the opportunity to do.

Let us consider another example. The sun rose this morning (I define rise to mean gradually appear over the horizon at some point in the morning); the sun rose yesterday; the sun rose the day before yesterday, and so on, for as long as I can recall. What does this, by itself, make the probability that the sun will rise tomorrow? The answer is that it doesn't say anything. One event does not make another event more or less probable unless we have some theory connecting the two events. So what we have to do is construct a theory. So, suppose that I have some theory, W_1, which states simply that the sun will rise each morning. I test this against the previously observed evidence, and it looks pretty good. I then make the prediction, on the basis of this theory. I say that if the theory W_1 is true then the probability of the sun rising tomorrow is greater than some number which is only a very tiny bit less than one (the best I can do on a frequentest account). Now, suppose that I have some colleague, who is smarter than I am, and who comes up with another theory W_2, which is derived from the classical general relativistic model of the solar system. Once again, he checks his theory against the prior observations that the sun has risen each morning, and it looks pretty good. But W_2 predicts that the sun will not rise tomorrow: there will be a total solar eclipse at dawn. So what is the probability that the sun will rise tomorrow? If we use theory W_1 we get one answer; if we use theory W_2 we get a very different answer. Both theories are in full accord with the previous evidence. So the probability $P(M \setminus W)$ does not depend on the evidence itself, but on which model we use to interpret the evidence.

Now, it may be objected that we have more evidence than just the observation that the sun has risen each morning to judge between theories W_1 and W_2, but this again misses the point: this is a failure of the analogy rather than the idea that the analogy is trying to illustrate. Hume is asking us to evaluate $P(M \setminus W)$ on a finite body of evidence; but that finite body is all that is available. In the case of the solar system, we do have more evidence available which contradicts theory W_1, but in the case of miracles we don't: both W_1 and W_2 are valid. Hume is postulating some theory W_1; this assumes that the universe is perfectly regular (as Hume understands regularity), and that the circumstances surrounding the event M are the same as the circumstances surrounding the events used to test the theory. This predicts that $P(M \setminus W_1)$ is infinitesimally small. But let us suppose instead that we have some theory W_2, which incorporates the idea that the circumstances surrounding M are different in some respect from those surrounding those events against which we have tested the theory. W_2 again passes the test of our general body of evidence, but predicts that $P(M \setminus W_2)$ is sufficiently high that the eyewitness evidence can push us into believing that the event occurred as described.

Let us suppose for the moment, to make this illustration simpler, that W_1 and W_2 are the only two possible theories that agree within experimental error with the observed evidence W (I will extend beyond this assumption below), and that they are mutually contradictory so can't both be true at the same time. We can then extend Bayes' theorem to cover this case. We write $W_1 \cup W_2$ as the combined probability of either W_1 or W_2 being correct. We can then write,

$$P(M \setminus (W_1 \cup W_2), W) = P(M \setminus W_1, W)P(W_1 \setminus W) + P(M \setminus W_2, W)P(W_2 \setminus W). \quad (718.1)$$

We need to estimate $P(W_1 \setminus W)$ and $P(W_2 \setminus W)$. I will write

$$P(W_1 \setminus W) = \frac{P(W \cap W_1)}{P(W \cap W_1) + P(W \cap W_2)}; \quad (718.2)$$

this is where I make the simplifying assumption that W_1 and W_2 are the only two possible theories. This assigns a number $P(W \cap W_1)$ which describes the intersection or correlation between the predictions of the theory W_1 and the observational data W (note the importance of experimental and theoretical uncertainty when computing this probability: it may be that

both models are consistent with the data, but one fits it better than the other). Of course, our total probability $P(W_1 \setminus W)$ not just depends on this number but also on how well the rival theories fit the data. Adding in the new evidence E and applying this to Bayes' theorem we find,

$$\begin{aligned} P(M \setminus E,(W_1 \cup W_2),W) =& \frac{P(M \setminus E,W_1)P(W \cap W_1)}{P(W \cap W_1)+P(W \cap W_2)} + \frac{P(M \setminus E,W_2)P(W \cap W_2)}{P(W \cap W_1)+P(W \cap W_2)} \\ =& \frac{P(E \setminus M,W_1)}{P(E \setminus W_1)} P(M \setminus W_1) \frac{P(W \cap W_1)}{P(W \cap W_1)+P(W \cap W_2)} + \\ & \frac{P(E \setminus M,W_2)}{P(E \setminus W_2)} P(M \setminus W_2) \frac{P(W \cap W_2)}{P(W \cap W_1)+P(W \cap W_2)}. \end{aligned} \tag{719.1}$$

The ratio $\frac{P(E \setminus M,W_1)}{P(E \setminus W_1)}$ describe how the eyewitnesses affect the evidence. This is the probability of getting that particular eyewitness testimony given the miracle occurring and our theory W_1 divided by the probability of that particular eyewitness testimony occurring regardless of whether or not the miracle happened (with everything else kept the same). This ratio is always greater than or equal to one, but how much greater than one depends on the reliability of the witnesses. If our witnesses are reliable, then they are considerably more likely to give their testimony if the miracle occurred than they would be to just make it up on the spur of the moment. If they were dishonest and (for example) stood to get financial gain from their testimony, then we might expect them to make the report in those circumstances regardless of whether the miracle occurred, and this ratio would be one.

By a similar process of reasoning we find,

$$\frac{P(M \setminus E,(W_1 \cup W_2),W)}{P(!M \setminus E,(W_1 \cup W_2),W)} = \frac{\frac{P(E \setminus M,W_1)}{P(E \setminus W_1)} P(M \setminus W_1) \frac{P(W \cap W_1)}{P(W \cap W_1)+P(W \cap W_2)} + \frac{P(E \setminus M,W_2)}{P(E \setminus W_2)} P(M \setminus W_2) \frac{P(W \cap W_2)}{P(W \cap W_1)+P(W \cap W_2)}}{\frac{P(E \setminus !M,W_1)}{P(E \setminus W_1)} P(!M \setminus W_1) \frac{P(W \cap W_1)}{P(W \cap W_1)+P(W \cap W_2)} + \frac{P(E \setminus !M,W_2)}{P(E \setminus W_2)} P(!M \setminus W_2) \frac{P(W \cap W_2)}{P(W \cap W_1)+P(W \cap W_2)}}. \tag{719.2}$$

If there are more possible theories than just W_1 and W_2, then we just have to add in the additional terms associated with these theories into the numerator and denominator of this equation, and replace the term $P(W \cap W_1)+P(W \cap W_2)$ with $P(W \cap W_1)+P(W \cap W_2)+P(W \cap W_3)+\ldots$. Since the equation is complicated enough already, and sufficient to make my point, I will stick with just the two theories (or classes of theories) and leave it to the reader to extend it in this way and make the minor modifications required for my argument below if he so desires.

Now W_1 represents some mechanistic world-view; Hume assumes that this is correct - Hume is erroneously setting $P(W_2 \setminus W)$ to zero. Even Hume's definition of miracle implies that it is a breaking of the laws of physics, which implies a mechanistic world view since in a theistic world view we don't usually discuss laws of physics as the most fundamental principle, and a miracle isn't strictly speaking a breaking of them. Now, $P(M \setminus W_1)$ is indeed very small, perhaps zero, and certainly small enough to overcome any eyewitness testimony beyond miraculous testimony. But this is not the whole story, because we must also consider the possibility that theory W_2, theism, is correct. In this case, we need some way of evaluating $P(M \setminus W_2)$; but, as we have seen, it need not be small. Indeed, as we have seen, that mankind falls short of goodness and cannot obtain it by his own efforts demands some divine intervention. Thus we expect that some miracle would occur at some point in time. Our probability $P(M \setminus W_2)$ thus depends mostly on how well the supposed miracle fits in with the overall theistic world-view. For something trivial, such as the branches of moving around to protect a man's dignity when he was having a bowel movement[10], we might assign this number to be quite small, since this isn't really a matter of any importance. On the

[10]Sahih Muslim 3005, Book 55, Hadith 93.

other hand, the resurrection of Christ purports to completely explain the Christian world-view: if Christianity is correct, then something like the resurrection had to have occurred. In this case, we can assign $P(M \setminus W_2)$ as a relatively high number: still small (we would not accept it without the eyewitness testimony) but not so small that the eyewitness testimony cannot affect our overall conclusion. This, of course, changes the formula completely: if our evidence is good, and the ratio $\frac{P(E \setminus M, W_2)}{P(E \setminus W_2)}$ very large and the ratio $\frac{P(E \setminus !M, W_2)}{P(E \setminus W_2)}$ very small, then as long as $P(W \cap W_2)$ is not significantly smaller than $P(W \cap W_1)$, then our evaluation of $\frac{P(M \setminus E, (W_1 \cup W_2))}{P(!M \setminus E, (W_1 \cup W_2))}$ will be a large number.

Hume's biggest mistake (in this passage) is that he assumes that theory W_1 is correct simply because it seems to be in agreement with the background evidence, ignoring that there may be other models in which the miracle is quite possible which are also in agreement with the background evidence. For example, suppose that Hume, Spinoza and Nietzsche got together to try to determine the probability of walking on water. They try it numerous times, and fail. After drying themselves off, they retire to dinner with Socrates, where they explain that walking on water must be impossible (for all practical considerations), and therefore the accounts that Jesus walks on water should be dismissed. But at this point Socrates interjects that they haven't proved that Jesus can't walk on water, but merely that Hume, Spinoza and Nietzsche were unable to walk on water on that occasion. Indeed, even they tried it out with everyone except Jesus, it would say nothing about whether Jesus would be able to do so. Hume retorts that there is no reason to assert that Jesus would be any different from them. Socrates replies that their evidence provides no reason to assert that Jesus is the same as them. And there is reason to say that Jesus is different, namely the testimony that he walked on water. Indeed, this is the whole point. Christians claim that Jesus is different: Jesus is God, and Hume is just some Scottish bloke. Nietzsche would probably comment that Socrates' argument is circular, which Socrates would happily admit, before Spinoza would make some mention of the laws of physics and how they show that people can't walk on water. And here Socrates would come in with his main point: that in phrasing the argument as they have, they have assumed that Jesus is the same as everyone else, bound to the standard laws of physics. They have used this assumption to dismiss the evidence that Jesus walked on water. Then they exclaim 'we have proved that Jesus was just a normal man, because there is no valid evidence that he walked on water.' This argument is obviously circular. Let us instead suppose that Jesus is God. Clearly he would then have the ability to walk on water if theism is correct and physics is just a description of God's usual actions in upholding the world (and by saying that Jesus is God, we are immediately conceding theism as well). Equally clearly, because other people are not God, they would lack this ability; thus the earlier experiment failed. But if Jesus could walk on water, and standard physics doesn't necessarily apply to him, then we have no good reason to dismiss the evidence concerning this event (assuming that the witnesses are reliable), thus proving that Jesus is, in fact God (or in some way associated with God). Obviously this argument is also circular. But why is it any worse than Hume's? It fits the rest of the evidence just as well; in fact it fits the evidence better because it both predicts that Hume can't walk on water and explains the evidence testifying that Jesus did. Of course, neither argument is good, because they are both circular: but this only means that Hume's whole approach to judging the evidence for one particular metaphysical theory by implicitly assuming a different metaphysical theory is flawed. Hume isn't just getting the wrong answer: he is asking the wrong question. At this point Nietzsche would mumble something about Ockham's razor: the simpler theory (the one which requires the fewest unproven axioms) is usually better. But Socrates would simply reply that we are not simply comparing the statements that 'Jesus is God, so what applies to Hume doesn't apply to Him' and 'Jesus is not God, so what applies to Hume does apply to him', but two entire world views, and there has been no discussion about whether the Christian theistic world-view is any more or less simple than the mechanistic, or how we count unproven axioms

when neither world-view is yet fully expounded, or indeed whether or not Ockham's razor is a useful guide in this circumstance.

I should also mention that although Hume claimed that his argument was inductive, it was not, because he sifted the evidence using a prior precondition (that happened to exclude the evidence that ran against his conclusion). Hume's argument was partly inductive, but partly deductive, since he used as a premise that the physical universe is a closed system, which (of course) excluded theism not by argument or by evidence but by presumption.

The weaknesses in Hume's argument have been constantly and repeatedly pointed out; it should not be necessary to do so again, but unfortunately he is still cited and used by atheist and deist writers without mention or discussion of these detractors. Indeed, most contemporary arguments about the impossibility of miracles are derived from Hume's, some expressed a little less clearly, others a little more, but all with the same basic flaws.

We have two theories, the materialistic and Christian theistic. We have already shown that all physical experiments to date have been consistent with Christian theism. Let us suppose without proof that materialism is equally consistent with experiment. Then (neglecting arguments from fine tuning), the only evidence that can distinguish between these two world views are miracles. What we must not do, given that materialism is not certain, is use a model which presupposes materialism to judge whether or not miracles occurred. But this is just what Hume requires to get his argument to work. ○

Objection:
Historians assume as part of their method that the world runs according to fixed laws; and they use this to assess supernatural claims. Any discussion of miracles involves supernatural claims. Therefore such accounts are inaccessible to historians. Therefore, a miracle cannot be proved by an appeal to history. However, the only evidence for miracles are historical sources. Therefore miracles cannot be proved to have occurred; and since there is no valid evidence, we should dismiss any accounts of the miraculous. •

Response:
There are two senses of history; 1) what actually occurred in the past; and 2) the reconstruction of the past made by historians. (2) depends on the various assumptions that a historian makes when evaluating the evidence. (1) is independent of those assumptions. When asking whether a miracle occurred, we are interested in sense (1) of history. As soon as historians bring in presuppositions, there may be a disconnect between the results of sense (2) and sense (1). When evaluating the historical evidence for miracles, historians ought not to presuppose the results of their inquiry. Any historian whose method presupposes the impossibility of miracles therefore cannot be trusted to evaluate the evidence without introducing bias. Instead, the historian ought to adopt a more neutral set of presuppositions, so that the conclusions are not determined by the method. A historian should separate the evidence: one one hand he obtains a probability of the event occurring from the reliability, number, bias, distance in time from the events they describe and so on of his sources (something that is largely the same in any world-view); on the other hand he can produce a probability based on how well this fits in with the overall picture or world-view implied by his underlying philosophical presuppositions. There is no need for the historian to combine these numbers to get a final answer, especially when the study overlaps with areas in which he is not an expert, such as philosophy or theology. Rather, the honest historian would give answers for each different possible world-view (*If this class of world-views are correct, then the probability of the event having occurred given the evidence of the historical sources is* x_1*; but if that class of world views is correct then the probability is* x_2*; and so on*), and let his

more philosophical aware reader combine these numbers to get a final answer. o

Objection:

> Nothing happens contrary to the order of nature, and that this order subsists without pause or interruption, eternal and unchangeable. ...
>
> All that God wills or resolves involves the conception of eternal truth and eternal necessity. The intelligence of God not being conceivable as distinct from his will, ... to say that God thinks or that God wills is to affirm one and the same thing. Consequently, the same necessity, in virtue of which it follows from the nature and perfection of God that he thinks a certain thing such as it is, this same necessity, I say, implies that God wills the thing such as it is. But as nothing is absolutely true save by divine decree alone, it is evident that the universal laws of nature are the very decrees of God, which result necessarily from the perfection of the Divine nature. If, therefore, anything happened in nature at large repugnant to its universal laws, this would be equally and necessarily repugnant to the decrees and intelligence of God; so that any one who maintained that God acted in opposition to the laws of nature would at the same time be forced to maintain that God acted in opposition to his proper nature, an idea than which nothing can be imagined more absurd. I might show the same thing, or strengthen what I have just said, by referring to the truth, that the power of nature is in fact the Divine Power; Divine Power is the very essence of God himself. But this I pass by for the present. Nothing, then, happens in nature which is in contradiction with its universal laws. ... Nature consequently always observes laws, although all of these are not known to us, which involve eternal truth and necessity, and thus preserves a fixed and immutable course. Nor will sound reason ever persuade us to ascribe a limited power and efficacy to nature, and to conceive its laws as operative in a certain restricted sense only, and not universally; for, since the power and efficacy of nature are the power and efficacy of God, and the laws of nature are the ordinances of God himself, we must needs believe that the power of nature is infinite, and its laws of such extent that they reach and pervade all that is comprehended by the divine intelligence. Were they not so, what else could be inferred than that God had made nature so impotent, and given it laws and statutes so barren, that he is forced frequently to intervene anew if he would have these laws continued, and the frame of things upheld in conformity with his wishes, – a doctrine as remote from reason as can well be conceived.
>
> From these premises, therefore, viz. that nothing happens in nature which does not follow from its laws; that these laws extend to all which enters into the divine mind; and, lastly, that nature proceeds in a fixed and changeless course; it follows most obviously that the word miracle can only be understood in relation to the opinions of mankind, and signifies nothing more than an event, a phenomenon, the cause of which cannot be explained by another familiar instance, or, in any case, which the narrator is unable to explain[11].

[11]Spinoza, Theologico-Political Treatise, Chapter 6.

•

Response:
Quantum physics clearly shows Spinoza to be wrong, because the claim *Nature consequently always observes laws, although all of these are not known to us, which involve eternal truth and necessity, and thus preserves a fixed and immutable course* is proved false by countless experiment: natural objects do not follow a fixed and immutable course. Indeed, the whole argument presupposes a naive mechanistic metaphysics, and this picture is inconsistent with our best experiments. This, of course, does not preclude God; as we have seen, an explanation for physics arising from God's continual intervention in the world both allows God's freedom and fits in well with what we have occurred. But this picture of metaphysics does not preclude miracles, but positively permits them. Since Spinoza's argument depends on his own notion of these laws, his argument falls with them.

As for the claim that miracles are simply natural events whose causes are not apparent, this again is disproved by modern physics: we know enough now to say that such events could not occur without special providence; and secondly the cause of the miracle is not unknown, and assigned to God out of desperation, but positively understood by both the witnesses and those directly involved be due to God.

Note also that the definition of miracle used by Spinoza, that miracles are a violation of the laws of nature and thus a violation of divine will since those laws are a complete expression of divine will, is different from the one adopted here. Both the miraculous and non-miraculous derive from the will of God, but the non-miraculous from general providence, where we work on the approximation that God is impartial to mankind, and the miraculous from special providence where we consider events where that approximation in invalid. It is therefore possible to distinguish a miracle from a non-miraculous event while acknowledging that both proceed from the will of God.

The notion that there are laws which are always followed assumes a mechanistic rather than theistic understanding of physics, and thus this objection can only apply to deism and not theism. ○

Objection:

> Miracles cannot make known to us the essence and existence of God, nor consequently his providence, these great truths being so much better illustrated and proclaimed by the regular and invariable order of nature....
>
> The existence of God not being obvious of itself, it must necessarily be inferred from ideas, the truth of which is so unquestionable that no power can be assigned or even imagined adequate to shake them. From the moment we conclude from these ideas that God exists, they ought to present themselves to the mind as beyond the sphere of doubt; for could we imagine that these notions could be changed by any power whatsoever, then should we doubt of their truth, and consequently of our conclusion as to the existence of God also, the effect of which would be that we should no longer feel certain of anything. And then we really know of nothing that agrees with nature or differs from it, save that which we have shown to agree with or to differ from these principles; wherefore, could we conceive that aught could happen in nature from any power (whatever this might be) which was repugnant to nature, this would also be repugnant to these primary notions, and so would have to be rejected as absurd; or else we should be forced to doubt of first notions

(as we have just said), and consequently of God and of all conceptions whatsoever. Miracles, therefore, conceiving these as events contravening the established order of nature, are so far from proving to us the existence of God, that they would actually lead us to call it in question, seeing that without them we can be absolutely certain of the existence of God, as we truly are when we know that all things in nature observe a definite and unchanging course.

But suppose it is said that a miracle is that which cannot be explained by natural causes; this may be understood in two ways: either that it has natural causes which cannot be investigated by the human understanding, or that it acknowledges no cause save God, or the will of God. But as all that happens, also happens by the sole will and power of God, it were then necessary to say that a miracle either owned natural causes, or if it did not, that it was something which by any cause; in other words, that it was something which it surpassed the human capacity to understand. But of anything in general, and of the particular thing in question, viz, the miracle, which surpasses our powers of comprehension, nothing whatever can be known. ... They therefore plainly trifle who when they do not know a thing, fall back upon the will of God - a most ridiculous way of professing or excusing ignorance. Moreover, whatever other inference may be drawn from miracles, nothing, at all events, can be concluded from them in regard to the existence of God; for, inasmuch as a miracle is a limited act, and never expresses more than a certain limited power, it is certain that we can never from such an effect infer the existence of a cause whose power is infinite; we could at the most conceive a cause, the power of which was relatively greater. ...

Let us conclude, therefore, that we can know nothing by miracles of the existence and providence of God; on the contrary, that these attributes are far better inferred from the regular and unchanging order of nature [12].

•

Response:

The exceptions of miracles do not override our knowledge of general providence, as long as they are sufficiently rare. Thus they do not undermine the arguments (for example from causality) which lead us to understand God's nature, his timelessness, immutability, simplicity and so on. Thus, to say that miracles undermine knowledge of God is false. However, reason alone cannot tell us everything about God, and miracles do reveal things concerning God's nature which are otherwise inaccessible to reason; for example, that God is not impartial to humanity's plight. Indeed, if the Christian view is correct, miracles are necessary for man to know enough about God to establish a relationship with Him. It is miracles which provide the evidence that allow us to affirm the superiority of theism over deism or pantheism; without the evidence of the miraculous we would be reduced to worshipping the pitiful and limited God that Spinoza believed in. Furthermore, as stated above, those who attribute miracles to God do not do so out of ignorance but because it is implied as the only explanation that correctly describes the context of the miracle. Additionally, although miracles are limited in time and space; that they override the general providence of God implies that they are nonetheless of immense significance and power, since (in principle at least) only God or a supernatural agent acting either on behalf of God or within the limits set by God for its nature can perform such an act.

[12]Spinoza, Theologico-Political Treatise, Chapter 6.

The charge that we should disbelieve miracle accounts because they contravene the laws of physics or are due to some presently unknown law of physics because the laws of physics are by definition universal is simply begging the question. The issue at stake is whether the laws of physics are universal, and miracles are part of the evidence brought forward to show that they are not. Thus we should not presuppose such universality when framing the question: the answer will depend more on our arbitrarily chosen axioms than on the actual evidence. ○

Objection:
No miracle in history is attested to by a sufficient number of honest men that it can be believed. •

Response:
This objection (by Hume) is just begging the question. Without examining examples (and one would have to examine all miraculous claims to prove this statement), the statement is simply an assertion without evidence. Indeed, Hume himself referenced the miraculous claims of the Jansenists (a Christian sect that had split from the Roman Catholic Church, but did not really accept Protestantism either), which provided the most controversial and most prominent miracle claims of his era. But it is important to note why Hume dismissed these accounts: not because he had a logical argument against miracles (he did not), but because he assumed before examining the evidence that miracles were inconsistent with the regular conjunction of causes. ○

Objection:
Humanity craves the miraculous, and will believe any absurd story. •

Response:
This is in part an over generalisation: not everyone believes any absurd story, and secondly commits the fallacy that because some people believe extraordinary claims which are absurd, it does not mean that every extraordinary claim that has been believed is in fact absurd. ○

Objection:
Miracles only occur among unsophisticated (and hence less reliable) peoples. Once a culture reaches a certain degree of civilisation, the miracle stories stop. Thus they are more a result of people's ignorance than anything else. •

Response:
This confuses correlation and causation. It states that because miracles are believed more in societies with a less degree of general education, then it is a lack of education that causes miracles to be believed. However, it may be that there is some other factor associated with the supposed reduction in the miraculous as societies become more 'sophisticated'; for example in more 'vulgar' societies it may be that faith is in general stronger, and if miracles depend on faith, then this would also explain the observation. The objection notes two trends that have occurred side by side, and posits some direct causation to explain this; this is a well known fallacy, since it neglects that there may be some other factor determining both trends, or it may be coincidental.

Also, I do not consider it proven that miracles are less common today than in the past. There are numerous contemporary accounts of the miraculous which largely pass by the notice of modern secular culture. (That statement is no less of an unsupported assertion than the original claim that miracles have stopped. However, my discussion for now is concerned with the general considerations. Examples of miraculous events, some more reliable than others, are common in the Christian literature. In my experience, most orthodox and devout Christians worldwide have experiences of miraculous testimony, either personally or through close friends. Not all of these testimonies are reliable or stand up to scrutiny, but many do.)

Thirdly, this takes a view of human civilisation which I think cannot be supported by history. Our ancestors were not idiots. Many among the Greeks and Romans, for example, were smarter and more knowledgeable about the world than the average men and women of our generation. But we have modern science. Certainly we do. But, one does not need modern science to understand that people do not in general walk on water, or people born blind suddenly receive their sight, or people lame for decades suddenly start leaping around and praising God. People knew that these were breaks in the normal order – the Christians as much as anyone – and the early Christians were often ridiculed for holding that such things happened. But again, one does not need a fully developed philosophy to know that if God exists, He is responsible for the natural order being what it is, and therefore has the ability to overrule it; while if God does not exist obviously He can't do so. The only advantage our modern science gives us is that it makes the one loophole possible in the ancient world considerably less plausible – a belief that such events are possible without God's (or some other supernatural agent's) intervention. As such, a more advanced science supports the claim that the supposedly observed event was miraculous, and therefore the work of God or some other supernatural agent. It is easy to call our ancestors barbarous, ignorant and gullible but that doesn't mean that they were any more so than we are today. The only difference between our gullibility and ignorance and that of our ancestors is that we find different sets of wrong ideas appealing. ○

Objection:
Miracles occur in all religions, implying different conclusions. These effects cancel each other out. •

Response:
All miracle claims count as evidence against atheism.

When judging a religion, or set of religious metaphysical principles, one should, of course, include miraculous evidence against that tradition as well as in support of it. Not every claim of a miracle is correct – I do not know anyone who says that they all are. However, we should judge each claim based on the reliability of its sources; for example for a miraculous healing we need to examine that the sickness was present, that it was cured, and that no other explanation for the cure could be found. When examining evidence for miracles, we need to consider both the numerous accounts of miracles in the present day and historical accounts, and the type of evidence in each case will be different. Just as the quality of evidence varies from one miracle to another, so does the quality of evidence for different religious traditions for miracles.

For example, the earliest written accounts we have for the miracles attributed to Muhammad were recorded some two hundred years after his death[13]; for the Buddha and miracles

[13]Muslims would object that the 'revealing' of the Qu'ran was a miracle that can be dated back to the time of Muhammad; I find it incomprehensible that anyone could think that composing a book of poetry

of the various polytheist traditions it is even longer. However, the testimonies of the miracles of Jesus and his apostles date back to sources contemporary with the eyewitnesses, and two of the Gospels and much of Acts are claimed to be eye-witness accounts themselves. The chains of individual witnesses between Muhammad and the later written records of the hadith might be reliable, but we know that when stories pass through several generations they can be embellished or distorted in ways that contemporary accounts passed on by a community and recorded by and evaluated by eyewitnesses cannot. Thus even though miracles are recorded in many religions, those accounts are not all of equal worth.

Furthermore, there is a difference between a miracle on which the key doctrines of a religion depend (and thus the miracle had to come first, and the religion later) and one which is superfluous to the religion's doctrines (and so could have been added on later). For example, the key miracles of Christianity (the incarnation, resurrection, and Pentecost) are the foundation of all Christian doctrine (remove them, and everything else collapses as well; therefore accounts of the resurrection had to come first and Christian doctrine later), while for many other religions they are irrelevant to the religious doctrine and can thus be discarded or added later without making any substantial change to the religion (therefore they could have been added relatively late in the religion's history).

Therefore while many religions have accounts of miracles, we can say that in some cases that the evidence is stronger than in others. That allows us to judge which set of reports (if any) to believe. For example, it is clear from the accounts from paganism, Islam, Buddhism, and others (including the later Gnostic texts related to Christianity) that it is quite common for people to add miraculous stories to some revered figure one hundred and fifty to two hundred years after the event. And advocates of these religions can't complain: the Muslim would have to concede that this process happened in Buddhism and Greek/Roman paganism, even if he denies if this process is the source of the miracle stories in the Hadith. However, having miracle stories in the eye witness accounts, or those of the same generation as the eye-witnesses is very rare: I am only aware of Christianity and Mormonism which claim contemporary records of miracles or have a vigorous ongoing tradition of miraculous events continuing to this day. The same process that gave rise to the Islamic or Buddhist accounts of miracles (if such accounts are false) cannot have happened (at least, not in the same way) in these religions, simply because the time frame is so much shorter. Thus, if the religions contradict each other, many can be eliminated from consideration because their evidence is considerably weaker than others.

Finally, believers in one religion need not suppose that all miraculous claims in other religions are false: lesser miracles could, for example, have been carried out by demons in order to lead people astray. I'm not saying that did happen, but in a Christian world-view it is a possibility. o

Objection:
Miracles are exceptionally improbable events. However, given enough samples, each exceptionally improbable event is likely to occur at least a few times. Therefore, accounts of the miraculous just occur in the natural order of things, and require no supernatural explanation. •

Response:
Suppose that some claimed miracle is a ten thousand to one chance. People pray, it happens, Hallelujah, glory be to God, sign me up for the book deal, and all that. After all, ten thousand to one sounds quite impressive. However, suppose that there are ten thousand

---

requires a miraculous explanation. Each culture and language has its own unmatched work of poetry; and many of them are of considerably greater literary quality than the Qu'ran.

similar opportunities occurring in different locations at about the same time. People pray, and in nine thousand, nine hundred and ninety nine cases the supposed miracle doesn't occur, while in one it does. This is, of course, roughly what we might expect. The nine thousand, nine hundred and ninety nine examples where it did not work are quietly forgotten, while the one case where it did happen is lauded throughout Christendom. This is, of course, nothing but a standard selection bias. Is this an adequate explanation of miraculous accounts?

In many cases, I would say that it is. Believing in miracles does not mean that one believes in all accounts of miracles; indeed it is good to be sceptical concerning any probabilistic miracle. For example a man who prays as he runs after a train, and happily discovers that his connection arrived at the station a few minutes late so he was able to catch it. Or the schoolboy who forgets to complete his homework on one occasion, prays, and that happens to be the occasion when the teacher also forgets to collect it, giving the boy time to rapidly complete it before the next class. Such things, even when repeated, are hardly enough to convince a sceptic, though they may prove edifying for the people involved. Of course, trains — even German trains — sometimes arrive late; teachers are not infallible. The miracle is the repeated coincidence with the prayers, though it is likely a mere coincidence. The mechanic in a factory who finds the engine of his truck filled with water: usually he would have to waste several hours at the machine shop having it fixed before it works again, but who instead declares to his colleagues that he has prayed, and the engine will work, to find it come to life. Again, such things are just coincidence. Or the honeymoon couple, setting out for a walk in a wooded canyon on a day forecast for rain, who pray for sunny weather. After emerging from the forest having been bathed in bright sunshine the whole time, they see two banks of dense cloud on the sky: a few hundred meters to the South of them, it rained continuously, a few hundred meters to the North of where they had been, there was also steady precipitation. But above them, stretching from East to West, a strip of unbroken, cloudless sky. But then, micro-climates are not unknown. Again, it is not so much the weather that is the marvel, but the coincidence that it occurred alongside the prayer. But, of course, we can also list the times of unanswered prayer, frustrated or delayed desires: the times when it did rain, when the engine was damaged further; these must also be counted. And that is the point of this objection: for every marvellously improbable event, there surely must be many more where the prayer failed which are not recorded or remembered.

We can divide claimed miracles into four distinct types (I do not discuss the reliability of the testimony that led to the claim here, as that is not the central point of the objection I am referring to). The first are those claims which are just plain wrong; where nothing occurred that is not explicable in terms of the usual symphony of cause and effect; there is nothing even unusual or coincidental about them. An example of this type of claim would be the recent Hindu statues which cried milk; which turns out to be merely a feature of the porous nature of the material the statues were made from, and their surrounding environment. Another example from a few decades ago was an outpouring in Canada, where members of a congregation 'miraculously' found themselves with gold teeth – curiously enough the same members who had recently received some dental surgery. We can also place into this class the numerous frauds and deceptions that have occurred throughout history, including Christian history. This type of story should, of course, be immediately discounted.

The second are those with a reasonably high probability of occurring naturally; for example the train mentioned above. These often build the faith of those to whom they occur, but cannot really be claimed as strong evidence for the miraculous: after all, such things happen without prayer as well as with prayer. All we are left with is an anecdote, and while that anecdote proves that trains are sometimes late when people pray, that is hardly enough to imply divine agency. To do that, one would need a large, controlled, ideally double blind sample, with large enough statistics to get a significant difference and the problem with such a study is that it is not possible. A miracle depends on an act of

God, and God is one crucial variable which is out of the control of the experimenter. The cause of the miracle is not the prayer, but God acting in response to the prayer, and God is free to act or not act as He pleases according to the wider circumstances. The one thing whose effects we are studying is the one thing we have no chance of isolating. I do not think that these stories can be used to justify miracles, even if some of the events are in practice miraculous. The odds for each individual event are too large to say that they were not likely to happen anyway. One in twenty; or one in a hundred occurrences are sufficiently common that we cannot really claim too much by it.

Then we get the third class of miracles; those events which are statistically highly improbable. Suppose that a skinny country pastor, moved by an image in a magazine of a young gang member convicted of murder, is moved to drive to New York city (where he has never been before) to console him and his family. One problem: the only thing he knows is that the family live somewhere in New York city. He has no address. A quick glance through the phone book fails: there are simply too many people in the city with the same name. So he drives around a bit, prays, and parks right outside the family's door (Wilkerson, 1963). There were about seven or eight million people in New York City at that time, perhaps two to three million households. But perhaps he knew the general area of the city he was aiming for, we can reduce the odds to just a few thousand or few tens of thousands to one. But then, maybe there were a thousand skinny pastors who drove to a thousand cities and nine hundred and ninety nine of them parked in completely the wrong part of town, and their stories were consequently never told. But the problem with this account is that it neglects that miracles of this nature tend to cluster around various people or places. So when that original pastor started seeing similar 'good fortune' repeatedly as he built up a ministry transforming gang members and drug addicts in New York, one can no longer use the defence of numbers. The first of these miracles (if it is a miracle) singles him out. And once we have focused on him, each subsequent one counts as evidence (although obviously any unanswered prayers in his ministry would count against the case). So if these stories are repeated around a single person (and the accounts taken to be reliable), then the cumulative case can be used as an argument for theism over mechanism, in the ways described below.

Then we come to the fourth class of miracles, those which are not merely highly improbable, but impossible in a materialistic world-view. I take the resurrection of Christ to be the clearest example of this; many contemporary stories of healings would fall into the same category (see (Keener, 2012) for a discussion of a small sample of modern miracle accounts). Now, these would on the face of it, if the accounts are reliable and it can indeed be truly proved that such things are impossible in a materialistic world-view, be firm proof that those forms of atheism which presuppose materialism, and also philosophies such as deism, are incorrect. Of course, those caveats are important. I will leave aside the issue of the reliability of the witnesses since I shall discuss that in more detail later. By what right can we say that such things are impossible without God (or some supernatural intervention)? Some claim that they may be due to an unknown law, because our knowledge is incomplete, but by now such claims should be dismissed. We do not yet know everything, but we know enough to fully rule out almost all possible theories which could describe physics, including any that might account for such stories being just natural events. Others might point to quantum indeterminacy as a way of making such events only merely highly improbable. But, as we have seen, this is not good enough. Quantum field theory is not a free for all where anything might happen – it still obeys various conservation laws, such as the conservation of electric charge and (most pertinently) of four-momentum (once we have fixed the gauge). Thus we can say that certain events are impossible without supernatural intervention. If these events occurred, then they contradict the view that the laws of physics are fundamental, which would undermine most forms of atheism and leave some form of theism (or some similar philosophy) as the only reasonable option. ◦

16.3. How to Assess the Evidence for Miracles

The problem we face is that it is impossible to estimate the probability that a miracle occurred except in terms of some metaphysical model. If, for example, we have the strongest possible evidence: numerous reliable and independent witnesses, detailed accounts both before and after the event happened, and a theory violated by the supposedly observed event that is as precise and as near to certain as it is possible to obtain, then whether or not we believe a miracle has taken place still depends on our prior assumptions. If we adopt some theist position, where God usually acts in a particular way but does not have to, and the supposed miracle is consistent with their general theology, then they will happily cite the evidence as a practical demonstration that that miracle occurred, and hence that miracles can occur, and thus put it forward in support of their own (or perhaps some similar) brand of theism. On the other hand, a materialist will look at the same evidence, and given their strong metaphysical prejudice against miracles, will conclude that, however unlikely it would seem, either the observations were faked, the witnesses not as reliable as we would naturally suppose, or that they were in some way deceived. If we then want to use the supposed occurrence of the miracle to judge whether or not theism or materialism is true, then we have descended into circularity.

I think that the problem is that we are asking the wrong question. The issue should not be 'Did this miracle occur?' but 'Are these historical records more consistent with one of the types of theism or one of the types of materialism?'

Of course, there are different degrees of evidence. If a supposed miracle is merely a unusual coincidence, then it isn't evidence against materialism. If the earliest records for the supposed miracle are several hundred years after the event, then it can easily be explained in terms of embellishment over time. If the supposed miracle is based on the testimony of just one person, with nothing physical to back it up, then that person can be dismissed as a liar. In all these cases the materialist has nothing to fear. The evidence is just as explicable if theism is false as it would be if theism were true. And, of course, many miracle accounts are like that. But, on the other hand, when there are numerous contemporary witnesses, backed up by hard physical evidence of the miracle, then it becomes harder to explain away. Of course, it is always still possible to do so. You can claim a conspiracy, or a magician's illusion, or that all the accounts and evidence was later faked. But such explanations always have an intrinsic improbability. It is not the miracle we are trying to account for, but the evidence for the miracle. The question is is this evidence more likely to have occurred if materialism is true or if theism is true. The theistic explanation for the existence of the evidence is straight-forward: God caused the miracle to occur, and the various records are what we would then expect to see. If the God in question is the sort of God who would display His power in that way and in those circumstances, then the probability of the records of the miracle existing given His existence would be quite high. Of course, we would have to be sure that the evidence is what we would expect if the miracle had occurred. For the materialist, on the other hand, the event could not have occurred. He has to explain why that evidence appeared then – and only then – when there was no event to cause it. The probability of that evidence occurring if some form of atheism exists is going to be small. And such explanations would also be possibilities in the theistic world view.

After all, when a historian sees numerous accounts of a battle in contemporary historians, and finds material remains of the battle, he doesn't then assume that it was a big conspiracy, for the simple reason that the probability of a conspiracy producing precisely the required evidence, with no surviving contrary accounts debunking it, is much less than the probability that the evidence exists because the battle actually happened. Hume's response that the prior probability of the miracle occurring is vanishingly small pre-supposes the truth of materialism, so is invalid in this calculation.

This is not, of course, to say that any such well-evidenced miracles exist. My purpose here is not to argue the case for any particular miracle, but to argue that, where such

evidence exists, we should take it seriously. There are plenty of accounts setting forth the evidence for particular miracles such as the resurrection of Christ (Wright, 2017), and I have no need to repeat those arguments again. When confronted with two different theories that each fit a particular set of data well, the good scientist will search for some evidence that can distinguish between them. Miracles are the evidence that, aside from any philosophical arguments, can allow us to distinguish between theism and atheism.

Let us not ask if the miracle occurred. Let us instead ask what are the possible circumstances which could have led to that particular array of documentary and archaeological evidence occurring, and then how plausible each of these circumstances are in each world-view.

So now we turn again to Bayes' theorem. We have some background evidence B, agreed on by both theists and materialists without dispute. $P(A_i \setminus B)$ is the probability that some particular materialistic world-view A_i is correct given the background evidence. $P(T_i \setminus B)$ is the probability that some form of theism is correct given this background evidence. We then have the records of this particular event, R. $P(R \setminus A_i, B)$ represents the probability that those records would be made if the materialistic world-view A_i were true, and, of course, still given our background evidence. The historical circumstances in which the miracle supposedly occurred might be part of our background evidence if known for certain, or can be incorporated into our probabilities if only known probabilistically. If we compare some materialistic theory A_j with some theistic theory T_k, then the evidence shifts the relative weight we give to these theories by

$$\frac{P(A_j \setminus R, B)}{P(T_k \setminus R, B)} = \frac{P(A_j \setminus B)}{P(T_k \setminus B)} \frac{P(R \setminus A_j, B)}{P(R \setminus T_k, B)}. \tag{731.1}$$

Of course, we then have to normalise the probabilities, which allows us (in principle) to find the absolute rather than relative probabilities of each theory,

$$\sum_i P(A_i \setminus R, B) + \sum_i P(T_i \setminus R, B) + \sum \ldots = 1, \tag{731.2}$$

where I have included the ... to cover any plausible world-views which are neither materialistic nor theistic. Once again, I emphasise that probability in the Bayesian interpretation is an account of how certain we are that the particular world-view is correct.

The probability that R could have occurred given B does not enter into these equations, although it is required in the earlier steps to derive the first equation. The mathematics suggests that we should consider the quantity $P(R \setminus B)$, however the probability of R should be independent of B, so it is good enough to use $P(R)$. Both R and B are collections of evidence, and it makes no sense to ask about the probability of one set of evidence occurring given another set of evidence occurring. A discussion of probabilities makes sense when comparing some evidence against a model, or some model against the evidence, but not when we compare one model against another or some evidence with other evidence. The details of the calculation require that $P(R)$ (or $P(R \setminus B)$ if my argument is unpersuasive) is non-zero. But this will be the case: some might say that there is some model of theism that is consistent with both R and B, and while that model remains possible, it remains possible that R could have occurred. But that misses the point: we have the records in front of us; they exist for certain: $P(R)$ is one, and therefore obviously non-zero. Equally, if we do not possess records describing the miracle, then $P(R)$ is zero and we cannot use this non-existent miracle to say anything because the derivation of (729.1) breaks down. We might, however, by a similar process use absence of evidence of miracles to count against various models of theism which positively predict that they ought to occur: we consider things like $P(!R \setminus B, A_k)$ and $P(!R \setminus T, A_k)$ where $!R$ indicates an absence of R. Clearly, $P(!R \setminus B, A_k)$ will be close to one (not exactly one; if we consider $P(R \setminus B, A_k)$ to be non-zero, then we can't make $(!R \setminus B, A_k)$ one), while the value of $P(!R \setminus B, T_k)$ will depend on the particular circumstances where the miracle did not occur and which type of theism we consider.

But even more importantly, the probability that the miracle occurred is not even required in the derivation of these equations. It is simply not something that need concern us when using miracles to judge between various forms of materialism and the various forms of theism. In this way, we can completely avoid any circular argument. So we are not asking (as Hume did) 'How likely is this event to have occurred given our materialist presuppositions and the historical records?' and then go on to look at the theism against materialism debate, but 'How likely are our materialistic presuppositions to be true given these historical records for this seemingly inexplicable event?' and also 'How central is this supposed miracle to this particular theistic world-view?'

Now, I have here made no assessment of what our original probabilities $P(T_k \setminus B)$ and $P(A_k \setminus B)$ are. This would, of course, depend on what we take B to be. Atheists would like to say that $P(T_k \setminus B)$ is so small that we may as well forget about this whole enterprise; the more stringent theists would say that $P(A_k \setminus B)$ is so small that this exercise is pointless (though it may be useful in distinguishing between different forms of theism). We can however, make some effort to determine the relative probabilities, or at least list some of the ingredients. Firstly, we have the requirement of self-consistency, but I have argued that classical theism is self consistent, and for the sake of argument I will suppose that materialism is as well. Then, how well do each theory match up with scientific experiment? And again, both world-views pass the test (I have demonstrated this for theism, and will suppose it for the sake of the argument for materialism, even though in practice, as I stated in earlier chapters, I am sceptical that mechanistic philosophy can be reconciled with observations at the quantum scale), at least as well as our current best theories of physics do. Both of these questions are objective: we can in principle (although it is a lot of work in practice) churn through each piece of data and do the calculation. Then we have to add in the arguments from fine tuning and causality; here we start to get a little more subjective: it depends on how plausible we find the multi-verse scenario, or how convincing we find the arguments from efficient and final causality. In my view, both of these arguments weigh the balance in favour of theism; indeed some say that the argument from causality when best presented (which may well be considerably better than I presented it in the earlier chapter), being a deductive argument, based on premises which are as empirically certain as anything we are likely to see, rules out mechanism altogether. On the other hand, there is no substantial evidence against theism[14]. So, in my view, the starting point is already firmly balanced towards the theist. However, the argument from miracles, when rightly presented, has the possibility to weigh the balance beyond any hope of doubt.

Objection:
The method proposed here suggests that we calculate the probability of some records appearing given some metaphysical model. So, we take the model, and, given a reasonable distribution of initial values, try to predict the probability that some historical record would have appeared. Now, let us take the (non-miraculous) example of Caesar's account of the Conquest of Gaul. The question we would be asking is to weigh up the likelihood of the various means by which the account we have could have come into being (i.e. as a authentic writing describing Caesar's own perspective on his battles, a late Roman or early medieval forgery, and so on). The method is asking us to take the basic facts of materialism: the laws of physics, Darwinian evolution, and so on, and from that asking us to calculate the probability of whether or not a man called Caesar would have written precisely that account down to the last stroke. Clearly, these odds will be so incredibly small that we ought not to bother with them. In any case, given that we cannot predict the existence of humanity, let alone Caesar or the Gauls from such fundamental premises, attempting to use

[14]People have argued against theism mainly by saying that it entails a self-contradiction; however I have not seen any valid argument of this sort. Those I have encountered have either been directed against a straw-man version of theism, or contain logical mistakes and misunderstandings.

this method cannot give any useful results. Thus, this method of trying to create an argument from miracles, whatever its merits or lack of them in principle, is in practice useless. •

Response:
Firstly, it is not necessary to take merely the laws of physics and biology as our only premises. There are, for example, certain facts which are agreed on by all the metaphysical models, such as of geography and some understanding of human nature, and these can safely be included as part of the basis of our calculations without introducing any bias towards one model or another. For example, in this case, what we believe about the basic details of the Roman Republic and its military structure are not dependent on any metaphysical model, and so can be assumed. As another example, when we discuss the miracles of Christianity, one of the key factors we ought to include in our deliberations are what the various groups of Jews in the early first century believed. Whether or not those beliefs were true is a question that affects the models we want to compare, so we can't either claim them to be true or false (or something in between) as part of our background knowledge B without introducing some bias; but what those beliefs were is unanimously agreed upon (up to some of the finer details) by well-informed people of any metaphysical bent. Thus, we can include the nature of those beliefs as part of the starting point of any calculations we make.

Secondly, it is not necessary in our analysis to set up a target so precise. For example, we might take R to be precisely Caesar's account; but it is better (given the questions we want to apply this method to) to be more general, and ask what is the likelihood of some Roman general writing some account of his military campaigns in Northern France at some point in history. Given the proximity of Rome and France, and the war-like nature of man, the odds of this occurring will clearly be something manageable.

Thirdly, for our purposes, there is no need to be precise in our calculation of the odds. Some reasonable estimate would be good enough (and all we can probably in practice hope to achieve).

Fourthly, we are interested in relative rather absolute probabilities. A lot of the improbability implied in the objection – the existence of Julius Caesar and the Gauls given knowledge of physics and biology, for example, is common to all the various metaphysical models and models of how the manuscripts of the book we possess came to be. This improbability will thus cancel out of our equation, and thus we need not bother computing it.

Fifthly, we already use this sort of argument even in the case of Caesar. Suppose that we have two models; model X being that Caesar conquered Gaul broadly as described in the book; model Y being that Caesar never ventured out of Italy beyond the Rubicon; the record R is the existence of the book and all we know about its acceptance and distribution in ancient Rome. Clearly the existence of the text is far more likely in model X than in model Y: it was written and received as a factual account, and there were too many contemporary witnesses for it to be accepted if it were a fictional political tract. Of course, that does not mean that Caesar did not embellish certain parts of his account – it was still written to bolster his political career, and we know from long experience that politicians rarely tell the full truth if they think that an exaggeration will boost their self-interest and they will be able to get away with it. But it could not have been wholly fictional. Thus the probability of $P(R \setminus Y, B)$ is considerably smaller than $P(R \setminus X, B)$; and it is for reasons such as this that historians prefer model X to model Y. ∘

So, when considering the records of a miracle, we have several things we need to consider. We should always ask how central the miracle is to a religion (which contributes to $P(R \setminus T_i, B)$). Given the nature of the religion, do we expect something like that to have occurred? Is the miracle a premise of the religion's doctrines, a conclusion of them, or wholly incidental to them? If it is a premise, then it is much more likely that the miracle came first and the

religious doctrine was developed around it – remove the miracle, and some other explanation needs to be found to explain the origin of the specific religion; if a conclusion then one could well imagine that it was invented as something that *must* have happened given the truth of the religion; if incidental then the miracle might be something added on to a set of religious doctrines and can just as easily be taken away. Another way of phrasing this is to ask how much of the religion is explained in terms of the miracle. For example, there are some (supposed) miracles which are wholly incidental to the doctrines of the religion. Get rid of these from the account, and the religion would be reasonably unaffected in its doctrine and practice. On the other hand, there are other (supposed) miracles which are central to the doctrine of the religion: remove these, and everything else collapses.

These central miracles are the best ones to consider when trying to weigh up the evidence for the religion. Firstly, because the most is at stake: if the evidence proves inconclusive, then you have most effectively undermined the religion; if it proves conclusive, then you have most effectively demonstrated the religion's correctness. Secondly, because the evidence for these tends to be most abundant. Thirdly because the religious belief does a better job of predicting that these will occur. Our method requires that we estimate the probability of the miracle occurring given the underlying metaphysical system; this probability will be much higher for central rather than incidental miracles.

There is a large difference in how we should weigh accounts of a miracle which is central to a religious tradition and one that is incidental to it. If the miracle story is central to the tradition, then it had to have originated with the religion, even if the first record of it is relatively late. You cannot found a religion and then thirty or fifty years later say, 'Oh, and here is one more thing which we forgot to mention before, but it is really the most important thing; everything else depends on this.' Clearly everything else does not depend on it, because everything else in the tradition was present without it. If everything else in the tradition does not even make sense unless it is an outworking of the supposed miracle, then the accounts of the miracle had to have come first and the tradition and dogma afterwards. When you preach a religion, you preach the most important parts of it first: its central miraculous claims; you then spend centuries trying to figure out some way of making the original prophet's utterances appear self-consistent, and this is when the theologies and doctrines and dogmas are created. Attempting to add a novel central claim to a religion well after the dogma has settled or the founder of the religion died would never be accepted, or at the very least it would be disputed and there would be some record of that dispute in the literature. The story had to originate before the religion had gained many devotees or reached a fixed form. On the other hand, if you have some additional incidental detail, of little importance, you can see how it could be added on later as the tradition developed. The central story of the religion is unchanged; no point of doctrine depends on it. People can believe that it might have have been overlooked when they first heard the religion. This is not to say that such a story *was* added as the religious traditions were passed down the generations; merely that it plausibly *could* have been, while a central miracle could not have been added to the tradition at a late date.

We should also ask about the quality of the witnesses to the supposed miracle. How many of them were there? Were they independent? Was the miracle unanimously accepted by the witnesses, and followers of the religion? Were any of the original witnesses converted on the basis of their witnessing of the miracle? How soon after the events was the testimony recorded? We generally give accounts written down in the lifetime of the original witnesses a higher weight than those written down several generations later: there is less chance of exaggeration, embellishment or fabrication, and more chance of the account not being refuted by other eyewitnesses the further away from the source the written records were. While oral tradition (particularly in ancient societies in certain circumstances) can be very reliable, it also nonetheless gives a good chance of introducing errors, particularly in points

which are not relevant to the core of the religion. All of this should be familiar to the historian.

However, one tool prominent in historical studies is not so relevant here: the question of bias. Historians generally give accounts biased in favour of some event less weight than those who are neutral or opposed to what is being recorded. For example, if we find that someone wrote, 'Pompey was the greatest general,' then we will take this testimony more seriously in our evaluation of Pompey's abilities if it were written by one of Pompey's enemies and less seriously if it were written by one of his flatterers. One can admire and respect Pompey's generalship even if one loathes the man himself and what he fought for. However, the case of a miracle is different: they are, by nature, events which turn heads. If you witness a miracle, you are likely to be convinced of the associated theology. Few people would say, 'I saw Elijah raise the widow's son from the dead, and am convinced that it was not a conjuring trick or fraud, but I still don't believe that there was anything special about the man.' You might claim that Elijah was sent from God, or perhaps from the devil, but you would not be able to claim that he was just your average bloke. Thus we expect the overwhelming majority, and perhaps a unanimous majority, of the witnesses to the supposed miracle to accept the underlying theology. Dismissing them because of their bias is not reasonable. Thus we expect to find the positive accounts of miracles almost exclusively contained in the accounts of the followers of that religion: to expect enemies to testify to them, perhaps baring a few snide comments about the gullibility of the believers, is not reasonable.

The time between the event occurring and the miracle being recorded is also of crucial importance. We know that if a story passes through a line of sequential witnesses, passed on from one person to another, then over several 'generations' the story can be embellished and distorted. An honoured man can thus 'acquire' miraculous powers over time. However, this process requires two things: time, and that the chain of witnesses should be individuals. It is difficult to falsely ascribe miraculous powers to someone while the eyewitnesses are in your audience. Perhaps he was like a modern magician; a master of deception who fooled everyone around him, although this doesn't really fit the stories we are interested in. But unless he fooled his own disciples, then those disciples will want to keep the story accurate. It is only after the passing of generations, as the first eye witnesses die out, that the miraculous elements can be added to the story.

Secondly, if the transmission is performed within communities, it is far harder to change or add new elements. If you pass on a story, not by whispering in someone's ear, but instead in the presence of a crowd, half of whom know the story at least as well as you, and all of whom consider accuracy of paramount importance, you cannot make major changes to the story without there being voices of dissension. The community memory acts as a means of self-correction in a way that is impossible in an individual chain of receivers. Of course, over time, minor details can change, and gradually add up until they make a major change, but again, this is a process which needs to take place over several generations; a matter of centuries rather than decades.

Thirdly, stories are less likely to be altered if they are repeated frequently. This is mainly a matter of memory; the speaker will remember the story better if he keeps it in the forefront of his mind. But equally, if the audience hears the story often, they will remember it better, be more likely to spot any inaccuracies, and be less likely to make a mistake when it comes to being their turn to pass it on.

Finally, multiple attestation is important: if several different people independently relate the same story, then what they have in common is likely to go back to the single source that informed each of those storytellers; the differences are likely to be due to the changes made in the time between the source and the eventual tellers of the story. These changes are uncorrelated, and there is no good reason why the same change should be made in two independent chains of transmission. Thus, from the similarities between the stories, the original story can be constructed.

An 'unintentional' addition of miracle stories to a religious tradition thus requires time and a specific set of circumstances. What about deliberate deception? This could only be done by an authoritative source and early in the history of the religion. Trying to convince a believing faithful of new stories, particularly if they change the content and doctrine of the religion, is immensely difficult, because they would see the evidence for the version they first heard as far stronger than for the novelty. Minor alterations around the margins of the religion could perhaps be tolerated, but at the centre of the religion? Not without creating a fuss which would be recorded in history. Thus any deliberate deception on major changes to the religion had to enter into the tradition at the start, from the first disciples, and be consistently repeated while the religion was preached to the first believers. How likely this is depends on the people involved, circumstances, and how much the religion is devoted to the truth. If they believe that they have an essential and necessary revelation from God, to change that in any way risks endangering souls, and runs against the very motivations and purposes of those initial apostles. The question is why would they lie, with so much at stake depending on the true message being heard? Of course, we should still look out for obvious charlatans, who preach for profit or personal power, or who were known for backing down when the going got tough. The more witnesses there were telling the same tale, the less likely it is that their story was faked. An individual, one may doubt. A crowd of five hundred all telling precisely the same story, particularly if they split up and were unable to conspire, is much harder to question. We cannot just say that it is plausible that the initial apostles could have lied: we have to look at the circumstances.

Now, it should also be clear that every account of a miracle which we possess provides evidence against atheism. This evidence might not be very strong, but it remains evidence nonetheless. Now given the vast quantity of miracle stories we have, on the face of it this seems to leave atheism in bad shape. Of course, we also have to balance against this the evidence of unanswered prayers: those occasions when we might expect a miracle were theism true, but nothing happened. Now obviously, in theistic systems such as Christianity, God is not obliged to answer any prayer. Thus an unanswered prayer is not proof against Christianity; but it does nonetheless count against it because the probability we would calculate given Christianity and the circumstances of the prayer that God would answer that prayer is almost certainly non-zero. How many unanswered prayers cancel out the effects of one actual miracle is difficult to say as a general rule. However, we might naively expect (if no significant numbers of true miracles have occurred) that if we do the calculation right that the few 'miracles' that occur by chance would be equivalent to the many unanswered prayers, leaving the relative probabilities of theism and materialism unaffected at the end of the calculation. In this way, this method need not lead to bias one way or the other if it is performed correctly.

So when assessing a miracle story, we should ask the following questions:

(1) Do the witness accounts of the supposed miracle require that it was an act of God? Is there no 'natural' explanation? For example, a man rising from the dead is impossible through natural means; the appearance and recital of a book of poetry, even superlative poetry, requires no supernatural explanation.
(2) How central is the miracle to the religious doctrine? Was the doctrine derived from the miracle, or the miracle story added to the doctrine?
(3) How likely is it for the particular religious beliefs to have emerged without the miracle in that particular culture?
(4) How close in time were the records of the miracle to the supposed event? How was it passed on between the event and the written record? How likely was it that the stories were embellished or forged?
(5) Were the miracle stories passed on in community or by individuals?
(6) How many independent witnesses were there to the miracle?

(7) Were the lives of the witnesses consistent with their testimony? Were they generally honest men?
(8) How much does the religion promote honesty? How much is at stake in the religion if false doctrine is allowed to mix with the truth?

16.4. Summary

Science is built upon two legs, theory and experiment, and all theoretical predictions demand to be tested through observation. In theology (whether theistic, atheistic, or something else), we can go a long way through philosophical argumentation, rightly built from sound principles. But this is limited; firstly because we cannot understand everything about God from reason alone, and secondly because we need to check our reasoning. In particular, reason tells us that God is there to be experienced; experience tells us that God is there to be understood rationally. It is not one or the other, but both working together. To know God, we need both special and general providence.

In this chapter, we have seen that, although many arguments have been made against the possibility of miracles, those arguments are based on materialistic premises. If we adopt theistic premises, however, and an understanding of what physics is built from theism, then we can distinguish the miraculous from the non-miraculous and miracles are to be expected. Arguments are only as good as the premises that they are built on, so the testimony for the miraculous is a direct challenge to atheism: if atheistic premises suggest that the miraculous is impossible, but there is good evidence that miracles have occurred, then those premises and thus atheism itself are false. Equally, there is a burden on theism – while a lack of evidence for the miraculous does not prove all brands of theism false (since theism only admits to the possibility of miracles, but does not predict when or where those miracles would occur), it certainly would be suggestive.

However, when evaluating miracle claims, we should not judge the evidence on the basis of our prior metaphysical commitments – whether the commitments that inspired Hume or Spinoza or the many others to declare that miracles are impossible, or the commitments that lead theists to believe that they occur. Rather, we should let the evidence we have, the records testifying to the claimed miraculous event, itself judge our theories and metaphysical presumptions.

The next step is obviously to fairly examine the evidence for miracle claims, and also for those situations where we might expect miracles to occur but they did not. I do not intend to do this here – that would take another book, and I have already written far too much in this one. But the evidence is out there: each religious tradition has already presented their case in numerous ways. I have suggested how such evidence might be evaluated. All I can ask, dear reader, is that you look with unbiased eyes.

CHAPTER 17

Conclusion

People address the question of the relationship between science and religion in various ways. Some say that they are non-overlapping magisteria; that they exist side by side without interfering. This is obviously partially correct – there are aspects of all religions about which science has nothing to say, and aspects of science where religions have nothing to contribute. But this is clearly not correct for the entirety of both areas of study; and this work should be an example of why. Others claim that the relationship has always been one of conflict, though this is equally obviously false, if nothing else because modern science arose in a religious context – the medieval European university. There are obvious points of interaction between specific doctrines of religions and various findings of science – evolution by natural selection, the age of the earth, and so on.

But, in my view, the most important place where science and religion meet is that both scientific theory and each religious or atheistic philosophy depend on various premises: to be precise metaphysical premises. Every rational inquiry depends on premises; this cannot be avoided. The self-consistency between premise and conclusion is no guarantee of truth. Inducing premises from conclusions can often be dangerous – it may be that more than one set of premises lead to the same conclusion. The best we can do is to rule out certain subsets of premises, because their conclusions are incompatible with observation.

In this work, I have suggested that Western philosophy took a serious wrong turn in the early modern period (though the mistake was first made before then). It was an understandable wrong turn, and perhaps a useful one, since it lead to the birth of mathematical physics. We needed to go through Newton to get to Feynman. This wrong turn was taken because it seemed that modern science suggested various things about the universe: that it was deterministic, that matter was constructed from small particles which could neither be created, destroyed, or changed in any way except in their locomotion or position, that observable properties could be simultaneously defined and uniquely correlate with a structure, and most importantly that particles were not inherently directed towards various goals or ends, but forced to obey external laws of motion. These were the assumptions that Galileo and Newton used as the basis of their physics, and which Descartes and Kant used as the basis of their philosophies. The empiricists, such as Hume, of course, took a different turn, and denied that we could know theoretical science or philosophy to any extent, and remained popular despite their key assumption being disproved even by Newton. Later philosophers built on these, and others of their generation, and consequently depended (in part) upon the same set of premises.

This is the world-view that has shaped our modern world; that has, in particular, spawned modern atheism, and also modern theological systems. It affects how we think about the world; it affects how we think about ethics; it affects how we view our place within our culture and our culture's place in the world, and it affects how we view history, politics and how we view science. Not how we do science, but how we answer the question what we consider scientific law to be. We do not need to be philosophers to feel the impact of this philosophy; it surrounds us in the whole culture, and is absorbed from school, university, and the media. Certainly few people bother to think through what they believe back to its premises. Many people label the philosophy as scientific, because it arose alongside Newtonian science, partly inspiring the science and partly inspired by it. However, this

claim is too bold even in the context of Newton's science, because they are not the only set of assumptions consistent with physical theory and with scientific observations. A distinction needs to be made between the physics of Newton (not to mention the twentieth century physics which overthrew Newton's theories) and the philosophy which inspired Newton and Galileo; too many people today confuse the two and label the philosophy as 'scientific' when it is anything but the result of science; nor even a necessary premise of science.

Incorrect premises almost certainly mean that the conclusions are wrong. Yet because the premises seem common sense to us, partly because they are drawn from Newton's physics which well describes the world at the scales which we are accustomed to, partly because everybody else seems to accept them, and partly because we were told that the alternatives were disproved at the start of the modern era, we rarely challenge them as rigorously as we ought to.

And challenge them we should. Physics has moved on, and quantum physics undermines almost every assumption that underpins the modern and post-modern ways of thinking. Most people today have accepted belief systems with foundations of sand. That the naive mechanism that motivated the founders of modern physics such as Galileo and Descartes is inconsistent with quantum field theory is, I think, obvious. The question is what should we put in its place: some modified form of mechanism to try to salvage as much of the enlightenment as we can, or something wholly different.

The rejection of Newtonian mechanics also means that we should reconsider the arguments which were made in the early modern period against the philosophy that preceded it. Classical mechanics was born from Christianity; its presuppositions were taken from a Christian world-view, and without theism it is not clear why the assumptions behind modern science should stand. However, it is widely believed that it has undermined theism in general. The various enlightenment philosophies both assume and are claimed to reject Christianity; much of what the modern world takes for granted and uses as its standard as it claims to be progressing to some universal nirvana was taken from Christianity, and yet it undermines that same Christianity more with each philosophical paper. The belief that modern Western secular society represents the culmination of every human civilisation, culture and morality is absurd: look at our society through any eyes except that of the Western academia, and you will see how frail, foundationless and inconsistent with the basic facts of humanity and physics almost all modern political, moral and philosophical discourse is. That contradiction between the roots of the culture and the direction in which it is heading means that there is a problem; and every problem cries out for a solution.

So we have two problems: experimental physics has turned against the philosophy that drove its early development, and enlightenment philosophy has turned against the world-view that provided the assumptions that it depends on and still needs, without supplying any alternative reason why those premises might be correct. I have proposed that both these problems have the same solution: to undo the enlightenment; to reinstate formal and final causes, now redefined in terms of contemporary physics but ultimately an expression of the same basic principles, and to accept a realism that acknowledges that both our understanding and observation are imprecise, but treats that uncertainty systematically and rationally. What I have proposed, I believe, combines the best of the medieval and modern traditions, and answers the problems inherent to both of them. I have shown that the premises behind Aristotelian philosophy can be induced from modern physical theory (in turn justified by experiment), needing only to be supplemented by a geometrical understanding of space and time and a theory of epistemology drawn from the modern post-Bayesian understanding of probability; that the premises behind classical theism can be deduced from Aristotelian metaphysics, and the premises behind modern physics can be deduced from the conclusions of classical theism.

Aristotle's philosophy is natural and self-consistent; it is based on the observation that there is both change and constancy in the universe, and that beings can both change in

some respects and remain constant in others. If this is granted, one is forced into a system of actuality and potentia (or something very similar to it), and it is from this and a few other principles which seem obvious that the Aristotelian philosophy, as perfected by the high medieval scholars of Western Europe, is derived. For those who understand it, it seems as though some variant of Aristotle's philosophy *must* be true, because the alternatives violate principles required for any understanding of nature. The later enlightenment philosophers did little but commit errors which the far more meticulous and careful scholastic philosophers had foreseen and deliberately constructed their system to avoid. Indeed, none of the problems or difficulties which plague modern philosophy affect the Aristotelian system, which resolves them or finds middle grounds in a self consistent and sometimes quite beautiful way.

Note that I am not, in this work, unlike the worst of the scholastic philosophers, assuming that Aristotle's metaphysics is correct and trying to fit physics to that. On the contrary, I assume that modern physics is correct, and search for the philosopher whose metaphysics comes closest to it, to use that philosopher as a starting point (not the ending point) for developing a philosophy of modern physics. That metaphysics, I judge, happens to be Aristotle's as it was developed in medieval Europe. Had physics turned out to be different, it could well have been Descartes, or Ockham, or one of the Chinese philosophers. But physics is what it is, and directs our hand. So why did Aristotle come closer to the truth than many of his more modern rivals? One reason is that modern philosophers tend to start by denying something as a first premise. The mechanists deny that substantial change is possible or that the universe could be indeterminate. The nominalists begin by denying the existence of substantial species. Kant began by denying a connection between the world of knowledge and the world of reality. The empiricists deny the possibility of knowing beyond what we directly observe; the philosophical relativists deny the possibility for objective truth. And, of course, Descartes' method of doubt started by denying everything except *cogito ergo sum*. Each of them start by making a presupposition that something isn't true; and when physical research shows that it does, in fact, happen to be true, that obviously puts them at a disadvantage. Aristotle's philosophy, however, starts from only acknowledging the existence of being, the possibility of change, both of which are self-evident, and that the universe is rational with conclusion following premise; substantial effect following substantial cause, which is necessary for any philosophical inquiry. As such, it is general enough to cope with a far broader range of physical theories than its rivals.

Is substantial change possible? Aristotle can cope with that, but since he does not require it to be necessary, he can also adapt to theories where it is not present. Are there universals? Aristotle's philosophy is a developed form of realism, but should there not be universals, then he can allow each species to only have a single member. Aristotle does not demand a deterministic physics (to within experimental precision), but nor does he contradict it. Aristotle's philosophy, being based only on what all rational studies of physics must presuppose, is sufficiently general to cope with anything. If physics were deterministic and one where substantial change impossible, then both mechanism and Aristotle's philosophy would be able to explain it; both would remain possibilities. But when that which the Enlightenment philosophers began by denying turns out to be true, only the more general theory survives: and therefore, we are led back to Aristotelian metaphysics.

Of course, the framework I have advocated in this work demands that we acknowledge the existence of God. But why is this such a burden? I have considered many of the modern objections to God, and shown that they fail, usually from the straw men of trying to force presumptions drawn from enlightenment philosophy into the classical tradition in ways that are clearly incompatible. I have instead suggested that the harmony of nature, the very structure of quantum field theory, is evidence for God. In this work, I have tried to suggest that belief in God in the contemporary world is neither irrational, nor unreasonable, nor

unsupported by evidence. Science is not incompatible with theism, but they fit together perfectly; the question we should be asking, in the light of the mid-twentieth century revolution in physics, is whether contemporary science is compatible with atheism. Modern atheism was born in the late nineteenth century from a view of physics which has now been firmly rejected. Instead, we should look at the direct evidence for God; not to dismiss miracles on the basis of perhaps unconsciously held assumptions, but to let the evidence dismiss our late nineteenth century superstitions of the necessity of human progress, and direct us to where reason is also guiding us back into the arms of God.

What is Physics? This is the question that is at the heart of the division between theists and atheists. It is not a question of science, but the findings of science can be used to judge between tentative answers. Does physics describe the outworking of blind external laws, that have an existence of their own and answer to no-one, as most people from the sixteenth to early twentieth centuries believed? Or is it a description of how God sustains the universe in the general course? Or is it some other option? The point is that our choice between these options is not blind; we can work out the consequences for each premise, and compare it against what we know of physical law. The first premise, that lies behind the atheism of the early twentieth century (such as that espoused by Bertrand Russell), leads to a mechanical universe, which we know to be false. The second premise, as I hope to have demonstrated in this work, when combined with classical theism, leads to a quantum field theory, and a class of physical theories which almost certainly contain the true theory of physics. Quantum physics seems hard to understand partly because people try to understand quantum mechanics rather than field theory, which is neither one philosophy nor the other, but mostly because people are so used to thinking with a Newtonian mindset that they forget that there are many other ways in which people can and have thought about the world.

What about the other possible premises beyond these two? Could they also imply quantum field theory? Perhaps. Could they be consistent with atheism? This is more troubling. That the material universe seems to be both indeterminate and directed, and source of that direction is external to matter since it cannot be predicted from even complete knowledge of the material universe (were such knowledge possible) implies an immaterial will; that that will is directed towards other objects and thus can comprehend them implies an immaterial intellect. We are getting very close to the God of theism.

Science or theism? Non-overlapping magisteria? Partially overlapping magisteria? No; science should be seen as a branch of theology, or its son, drawing its premises from a theistic metaphysics. The two necessarily separated for a time because of the path down which experiment led physics; but now physics is reaching its destination, and we find that it is returning back to its Father.

I started this work as a quest for the truth; well, I am not there yet. But hopefully, this time, our footsteps are at least pointing in the right direction.

Bibliography

I. Aitchison and A. Hey. *Gauge Theories in Particle Physics: A Practical Introduction, Volume 1: From Relativistic Quantum Mechanics to QED.* CRC Press, 2012.

W. Altmannshofer, W. A. Bardeen, M. Bauer, M. Carena, and J. D. Lykken. Light Dark Matter, Naturalness, and the Radiative Origin of the Electroweak Scale. *JHEP*, 01:032, 2015. doi: 10.1007/JHEP01(2015)032.

T. Aquinas. *Commentary on the Sentences.* 1256.

T. Aquinas. *Summa Contra Gentiles.* 1264.

T. Aquinas. *Summa Theologica.* 1274.

T. Aquinas. *Commentary on the Metaphysics (Metaphysicorum expositio).* c. 1272.

T. Aquinas and T. McDermott. *Selected Philosophical writings.* Oxford, 2008. Along with from Aquinas' major works, one should read as preparation for them *De ente et essentia*, *De principiis naturae*, and his commentaries on Aristotle, particularly *Posteriora Analytica*, *Physica*, *De coelo et mundo*, *Ethica*, *De generatione et corruptione* and *Metaphysica.* Translations of most of his works are available online.

Aristotle. *Metaphysics.* c. 330BC.

Aristotle and J. Barnes. *The Complete works of Aristotle.* Princeton, 2014.

P. Atkins. *The Laws of Thermodynamics: A Very Short Introduction.* Oxford, 2010.

A. Atland and B. Simons. *Condensed Matter Field Theory.* Cambridge, 2010.

M. Augros. *Who Designed the Designer?* Ignatius, 2015.

Augustine of Hippo. *On The Literal Interpretation of Genesis.* 404-420.

F. Bacon. *The New Oragon.* 1604. Translated by Jonathan Bennett.

J. Baggott. *Farewell to Reality: How Fairytale Physics Betrays the Search for Scientific Truth.* Constable, 2003.

L. Barnes. The Fine-Tuning of the Universe for Intelligent Life. *Publ. Astron. Soc. Austral.*, 29:529, 2012. doi: 10.1071/AS12015.

L. Barnes, G. Lewis, and B. Schmidt. *A Fortunate Universe: Life in a Finely Tuned Cosmos.* Cambridge, 2016.

L. A. Barnes. Fine-Tuning in the Context of Bayesian Theory Testing. *ArXiv e-prints*, July 2017.

J. Barrow and F. Tipler. *The Anthropic Cosmological Principle.* Oxford, 1986.

R. Baschab. *A manual of neo-scholastic philosophy.* B. Herder, St. Louis, 1937.

J. Bentham. *An Introduction to the Principles of Morals and Legislation.* Oxford, 1823.

R. Bhaskar. *A Realist Theory of Science.* London, 1975.

J. Binney and D. Skinner. *The Physics of Quantum Mechanics.* Oxford, 2013.

J. Bjorken and S. Drell. *Relativistic Quantum Mechanics.* McGraw-Hill, 1964.

J. Bjorken and S. Drell. *Relativistic Quantum Fields.* McGraw-Hill, 1965.

N. Bostrom. *Anthropic Bias: Observation Selection Effects in Science and Philosophy.* Routledge, 2010.

R. Boyle. *Of the Excellency and Grounds Of the Corpuscular or Mechanical Philosophy.* 1674. In my citations of this work, I updated the spelling and punctuation.

W. Briggs. *Uncertainty: the Soul of Probability.* Springer, 2016.

C. P. Burgess. Introduction to Effective Field Theory. *Ann. Rev. Nucl. Part. Sci.*, 57: 329–362, 2007. doi: 10.1146/annurev.nucl.56.080805.140508.

E. A. Burtt. *The Metaphysical Foundations of Modern Science.* Dover Publications Inc, 1931.

B. Carter. Large number coincidences and the anthropic principle in cosmology. In *Confrontation of Cosmological Theories with Data*, page 291. Longair, 1974.

G. Chesterton. *St. Thomas Aquinas.* Hodder and Stoughton, 1933.

W. L. Craig. *Time and Eternity.* Crossway, 2001.

P. Davies. *The Goldilocks Enigma.* Penguin, 2006.

R. Descartes. *A Discourse on Method.* Read Books (2008), 1637. ISBN 9781443733748.

P. Devine. *Natural Law Ethics.* Greenwood Press, 2000.

R. A. d'Inverno. *Introducing Einstein's Relativity.* Oxford University Press, 1992.

P. Dirac. *Lectures on Quantum Mechanics.* Belfer Graduate School of Science, 1964.

J. E. Dolezal. *God without Parts.* Pickwick Publications, 2011.

P. Draper. Pain and pleasure: An evidential problem for theists. *Noûs*, 23:331 350.

A. Drezet. Lorentz-invariant, retrocausal, and deterministic hidden variables. *Found. Phys.*, 49(10):1166–1199, 2019. doi: 10.1007/s10701-019-00297-5.

A. Einstein. *Relativity: The Special and the General Theory.* Methuen & Co. Ltd, London, 1920.

A. Fagothey. *Right And Reason: Ethics in Theory and Practice.* C.V. Mosby, 1959.

E. Fermi. *Thermodynamics.* Dover, 1936.

E. Feser. *Aquinas: A Beginner's Guide.* Oneworld, 2009.

E. Feser. Teleology: A shoppers guide. *Philosophica Christi*, 12:142, 2010.

E. Feser. The medieval principle of motion and the modern principle of inertia. *Proceedings of the Society for Medieval Logic and Metaphysics*, 10:4, 2012.

E. Feser. *Scholastic Metaphysics: A Contemporary Introduction.* Editiones Scholasticae, 2014.

E. Feser. *Neo-Scholastic Essays.* St. Augustine's Press, 2015.

E. Feser. *Five Proofs for the Existence of God.* Ignatius Press, 2017.

E. Feser. *Aristotles Revenge: The Metaphysical Foundations of Physical and Biological Science.* Editiones Scholasticae, 2019.

E. Feser and J. Bessette. *By Man Shall His Blood Be Shed: A Catholic Defense of Capital Punishment.* Ignatius Press, 2017.

R. Feynman. *The Character of Physical Law.* BBC, 1965.

R. Feynman and A. R. Hibbs. *Quantum Mechanics and Path Integrals.* McGraw-Hill, 1965.

A. French and M. Ebison. *Introduction to Classical Mechanics.* Springer, 2013.

G. Galilei. *The Assayer.* Rome, 1623.

G. Galilei. *Dialogue Concerning the Two Chief World Systems.* Florence, 1632.

G. Galilei. *Discourses and Mathematical Demonstrations Relating to Two New Sciences.* Leiden, 1638.

R. Garrigou-Lagrange. *Reality: A Synthesis of Thomistic Thought.* St. Louis, Herder, 1950.

N. Geisler. *Thomas Aquinas: An Evangelical Appraisal.* Wipf & Stock, 2003.

E. Gilson. *The Christian Philosophy of Thomas Aquinas.* University of Notre Dame Press, 1994.

E. Gilson. *Philosophy of St. Thomas Aquinas.* Kessinger, 2003.

N. Gisin. On the Impossibility of Covariant Nonlocal "hidden" Variables in Quantum Physics. *Physical Review A*, 83:020102–1/2, 2011.

J. Glover, editor. *Utilitarianism and its critics.* MacMillan, 1990.

Goldstein, Safko, and Poole. *Classical Mechanics.* Pearson, 2001.

A. Gomez-Lobo. *Morality and the Human Goods: An Introduction to Natural Law Ethics.* Georgetown University Press, 1999.

E. Grant. *Science and Religion, 400 B.C. to A.D. 1550: From Aristotle to Copernicus.* Johns Hopkins University Press, 2006.

E. Grant. *God and Reason in the Middle Ages.* Cambridge University Press, 2010.

D. Griffiths. *Introduction to Quantum Mechanics.* Cambridge, 2016.
D. J. Griffiths. *Introduction to Electrodynamics.* Pearson, 1998.
W. Grudem. *Systematic Theology.* IVP, 1994.
J. Hannam. *God's Philosophers: How the Medieval World Laid the Foundations of Modern Science.* Icon Books Ltd, 2010.
S. Hawking and R. Penrose. The singularities of gravitational collapse and cosmology. *Proceedings of the Royal Society London*, A314:529–548, 1970.
W. Heisenberg. *Physics and Philosophy.* Allen & Unwin, 1958.
P. Hoenen. *Cosmologia.* Gregorian University Press, 1956.
D. Hume. A treatise on human nature, 1739. Book III, Part I, section I.
D. Hume. *An Inquiry into Human Understanding.* Millar, London, 1748.
T. Irwin. *Plato's Moral Theory.* 1979.
T. Irwin. *Plato's Ethics.* 1995.
C. B. J. and R. M. J. *Nature*, 278:605, 1979.
J. D. Jackson. *Classical Electrodynamics.* Wiley, 1998.
M. Jammer. *The Philosophy of Quantum Mechanics: Interpretations of Quantum Mechanics in Historical Perspectives.* Wiley-Blackwell, 1974.
W. Jaworski. *Structure and the Metaphysics of Mind: How Hylomorphism Solves the Mind-Body Problem.* Oxford, 2016.
I. Kant. What is enlightenment, 1784.
I. Kant. Grounding for the metaphysics of morals, 1785.
C. S. Keener. *Miracles:The Credibility of the New Testament Accounts.* Baker Academic, 2012.
A. D. Kennedy. Algorithms for dynamical fermions. 2006.
A. Kenny. *The Five Ways.* Schocken Books, 1969.
F. Kerr. *Aquinas: A very Short Introduction.* Oxford, 2009.
G. Klima. *John Buridan.* Oxford, 2008.
P. Kreeft. *Summa of The Summa.* Ignatius, 1990.
L. D. Landau and E. M. Lifshitz. *Mechanics: Volume 1.* Butterworth-Heinemann, 1976.
B. Lawn. *The Rise and Decline of the Scholastic "Quaestio Disputata".* Brill, 1993.
J. Leslie. *Universes.* Routledge, 2002.
M. Leunissen, editor. *Aristotle's Physics: A Critical Guide.* Cambridge, 2015.
H. Leutwyler. On the foundations of chiral perturbation theory. *Annals Phys.*, 235:165–203, 1994. doi: 10.1006/aphy.1994.1094.
D. Lewis. *On the Plurality of Worlds.* Blackwell, 1986.
D. C. Lindberg. *The Beginnings of Western Science.* University of Chicago Press, 2008.
M. Loux. *Metaphysics: A Contemporary Introduction.* Routledge, 2006.
E. J. Lowe. *A Survey of Metaphysics.* Oxford, 2002.
B. S. M. and K. A. *Physical Review D*, 76:045002, 2007.
A. MacIntyre. *Dependent Rational Animals.* Open Court, 1999.
A. MacIntyre. *Ethics and Politics: Selected Essays.* Cambridge, 2006.
A. MacIntyre. *After Virtue.* University of Notre Dame Press, 2007.
A. MacIntyre. *Whose Justice? - Which Rationality?* Bloomsbury, 2013.
A. MacIntyre. *Ethics in the Conflicts of Modernity.* Cambridge, 2016.
G. Maclean. *Bertrand Russell's Bundle Theory of Particulars.* Bloomsbury, 2014. I have cited one version of the bundle theory of metaphysics here, in which essence is identified with a bundle of properties; the idea itself goes back to Hume or perhaps beyond him, and is popular among realist empiricists. Any good contemporary overview of metaphysics, such as (Loux, 2006) will discuss the various flavours of it.
J. Maritain. *St. Thomas Aquinas.* Meridian, 1962.
T. Maudlin. *Quantum Non-Locality and Relativity.* Blackwell, 1994.
A. McGrath. *A Scientific Theology.* T&T Clark, 2001.

R. McInerny. *St. Thomas Aquinas.* Notre Dame: University of Notre Dame Press, 1982.
D.-J. Mercier. *A Manual of Modern Scholastic Philosophy (2 Volumes).* Louvain, 1900.
J. S. Mill. *Utilitarianism.* Parker, Son & Bourn, 1863.
G. E. Moore. *Principia Ethica.* Cambridge, 1903.
S. Mumford. *Metaphysics: A Very Short Introduction.* Oxford, 2012.
I. Newton. *Philosophiae Naturalis Principia Mathematica*, chapter General Scholium. 1713.
I. Newton. *General Scholium.* Pepys, London, 1726. Translated by Andrew Motte.
I. Newton. A short scheme of the true religion. In H. McLachlan, editor, *Theological Writings.* Liverpool University Press, 1950.
D. Oderberg. *Applied Ethics: A Non-Consequentialist Approach.* Wiley, 2000.
D. Oderberg. *Real Essentialism.* Routledge, 2008.
D. Oderberg. Whatever is changing is being changed by something else: A reappraisal of premise one of the first way. In J. Cottingham and P. Hacker, editors, *Mind, Method and Morality: Essays in Honour of Anthony Kenny.* Oxford, 2010.
D. Oderberg and T. Chappell, editors. *Human Values: New Essays on Ethics and Natural Law.* AIAA, 2008.
C. Osborne. *Presocratic Philosophy: A Very Short Introduction.* Oxford, 2004.
T. Paine. *The Rights of Man.* Jordan, 1791.
Pallen and Matzke. From the origin of species to the origin of bacterial flagella. *Nature Reviews Microbiology*, 4(10):784–790, 2006.
R. Penrose. *The Emperor's New Mind.* Oxford University Press, 1999.
M. Peskin and D. Schroeder. *An Introduction To Quantum Field Theory.* Westview, 1995.
A. Plantinga. *The Ontological Argument from St. Anselm to Contemporary Philosophers.* Doubleday, 1965.
A. Plantinga. *The Nature of Necessity.* Oxford, 1979.
A. Plantinga. *Where the Conflict Really Lies.* Oxford, 2011.
Plato. *Πολιτεια (The Republic)*, chapter VII. Athens, 380 BC.
J. Polkinghorne. *Quantum Physics and Theology: An Unexpected Kinship.* SPCK, 2007.
A. Pruss. *The Principle of Sufficient Reason: A Reassessment.* Cambridge, 2010.
A. Rae. *Quantum Physics: Illusion or Reality?* Cambridge, 1986.
M. Rees. *Just Six Numbers:The Deep Forces That Shape the Universe.* W&N, 1999.
F. Reif. *Fundamentals of Statistical and Thermal Physics.* McGraw-Hill, 1965.
C. E. Rice. *50 Questions on the Natural Law: What It Is and Why We Need It.* Ignatius, 1999.
C. Rovelli. Aristotle's physics: A physicist's look. *Journal of the American Philosophical Association*, 1:22, 2015.
W. Rowe. *The Evidential Argument from Evil: A Second Look*, page 262. Indiana University Press, 1996.
J. Rowling. *Harry Potter and the Half Blood Prince.* Bloomsbury, 2005.
L. H. Ryder. *Quantum Field Theory.* Cambridge, 1996.
J. Sachs. *Aristotle's Physics: A Guided Study.* Rutgers, 1995.
P. Sanlon. *Simply God.* IVP, 2014.
S. Saunders, J. Barrett, A. Kent, and D. Wallace, editors. *Many Worlds?: Everett, Quantum Theory, & Reality.* Oxford, 2012.
A. Schakel. *Boulevard Of Broken Symmetries: Effective Field Theories Of Condensed Matter.* World Scientific, 2008.
V. Schieber, T. D. Conway, D. M. McCarthy, and H. P. CSJ. *Where Justice and Mercy Meet: Catholic Opposition to the Death Penalty.* Liturgical Press, 2013.
Schleiermacher. *The Christian Faith.* Reimer, Berlin, 1884.
D. V. Schroeder. *An Introduction to Thermal Physics.* Addison Wesley, 1999.
R. Shankar. Effective Field Theory in Condensed Matter Physics. *eprint arXiv:cond-mat/9703210*, Mar. 1997.

I. Silva. Werner heisenberg and thomas aquinas on natural indeterminism. *New Blackfriars*, 94(1054):635–653, 2013. ISSN 1741-2005. doi: 10.1111/nbfr.12036. URL `http://dx.doi.org/10.1111/nbfr.12036`.

P. Singer. *Practical Ethics*. Cambridge, 1979.

L. Sklar. *Space, time, and spacetime*. University of California Press, 1974.

V. E. Smith. *Philosophical Physics*. Harper & Brothers, 1950.

W. Smith. *The Quantum Enigma: Finding the Hidden Key*. Angelico Press/Sophia Perennis, 2005.

R. J. Spitzer. *New Proofs for the existence of God*. William B Eerdmans, 2010.

V. Stenger. *The Fallacy of Fine-Tuning*. Prometheus, 2011.

K. Stowe. *An Introduction to Thermodynamics and Statistical Mechanics*. Cambridge, 2014.

P. Strange. *Relativistic Quantum Mechanics: With Applications in Condensed Matter and Atomic Physics*. Cambridge, 2008.

W. Struyve. Pilot-wave approaches to quantum field theory. *J. Phys. Conf. Ser.*, 306: 012047, 2011. doi: 10.1088/1742-6596/306/1/012047.

E. Stump. *Aquinas*. Routledge, 2005.

D. Sullivan. *An introduction to philosophy : the perennial principles of the classical realist tradition*. Bruce, Milwaukee, 1957.

E. F. Taylor and J. A. Wheeler. *Spacetime Physics: Introduction to Special Relativity*. W. H. Freeman & Company, 1992.

J. Thijssen and J. Zupko. *The Metaphysics and Natural Philosophy of John Buridan*. Brill, 2001.

J. Townsend. *A Modern Approach to Quantum Mechanics*. University Science Books, 2012.

J. Unwin. *Sex and Culture*. Oxford, 1934.

H. Veatch. *Human Rights: Fact or Fancy?* LSU Press, 2007.

L. Velecky. *Aquinas' five arguments in the Summa Theologica 1a 2,3*. Peeters, 1994.

R. M. Wald. *General Relativity*. University of Chicago Press, 1984.

R. M. Wald. *Space, time, and gravity: the theory of the big bang and black holes*. University of Chicago, 1992.

R. M. Wald. The Formulation of Quantum Field Theory in Curved Spacetime. *Einstein Stud.*, 14:439–449, 2018. doi: 10.1007/978-1-4939-7708-6_15.

W. A. Wallace. *The Modeling of Nature: Philosophy of Science and the Philosophy of Nature in Synthesis*. The Catholic University of America Press, 1996.

R. Waterfield. *The First Philosophers*. Oxford, 2000.

S. Weinberg. *The Quantum Theory of Fields*. Cambridge, 1995.

D. Wilkerson. *The Cross and the Switchblade*. Pyramid, 1963.

D. Williams. On the elements of being. *Review of Metaphysics*, 7:3, 1953. Trope nominalism identifies tropes, or properties, as an additional layer on top of being; since Williams' initial work it has become a major stream of nominalism.

J. Wippel. *The Metaphysical Thought of Thomas Aquinas*. The Catholic University of America Press, 2000.

A. Wood. *Kantian Ethics*. Cambridge, 2008.

N. T. Wright. *The Resurrection of the Son of God*. SPCK, 2017.

Index

W-Boson, 347, 673
Z-Boson, 347, 673
γ-matrix, 657
ψ-epistemic, 634
ψ-ontic, 634
$\vec{\nabla}$, 212

A-theory of time, 239
Absolute space, 225
Accelerating frame, 243
Accident, 99, 370
Accidentally ordered series, 402, 438
Accuracy, 78
Actuality, 92, 114, 400
Aether, 171, 206, 217, 232
Affine structure, 231
Agape, 590, 594
Alchemy, 148, 153
Ampére, 213
Amplitude, 20, 50, 215, 309, 633, 636, 675, 683, 702
Animal life, 582
Animal soul, 111
Annihilation operator, 329, 358, 364, 626, 648
Anomaly, 674
Anthropic principle, 689
Anti-commutator, 331, 637, 647
Anti-matter, 324, 332
Anti-particle, 379, 649
Aquinas, 395, 396
Argument from design, 416
Aristotelian dualism, 186
Aristotelian metaphysics, 93, 285
Aristotelian physics, 127, 136
Aristotelian premises, 93
Aristotle, 11, 42, 127, 174, 180, 186, 194, 221, 287, 536, 559
Artificial, 100
Artificial motion, 114
Artificial purpose, 563
Astrology, 148, 177
Astronomy, 176
Atheism, 24, 260, 388
Atomism, 196
Axioms, 229
Axioms of Probability, 44
Axioms of Science, 54

B-theory of time, 239
Bacon, Francis, 155
Bacon, Roger, 154
Bayes' theorem, 616, 696, 710, 714, 727
Bayesian interpretation of probability, 47
Being, 94, 429
Being ethics, 496
Bell's inequalities, 650
Bell's theorem, 650, 668
Big bang, 250
Biological species, 359
Boltzmann, 293
Boson, 322, 331, 333, 637, 655, 658, 664
Boyle, 163, 169
Bradwardine, 144, 169, 371
Brute fact, 68, 405
Buridan, 146, 169

Cabibbo Kobayashi–Maskawa matrix, 349, 675
Calculus, 147, 201
Canonical quantisation, 328
Capabilities, 560
Cartesian doubt, 184
Cartesian dualism, 184
Categorical imperative, 525, 545
Causal physics, 142
Causality, 115, 196, 238, 264, 268
Change, 90, 114, 125
Character, 531
Chinese and Indian science, 134
Chiral gauge theory, 349, 671
Christoffel symbol, 246
Classical interpretation of probability, 44
Classical limit, 317
Classical mechanics, 14, 203, 309, 666
Classical science, 134
Classical theism, 87, 387, 625
Cluster decomposition, 364
Colour, 347
Communication, 590
Commutator, 312, 331, 637, 647
Complete, 44
Complex conjugate, 296, 309
Complex numbers, 296, 309
Composite, 109, 449
Conceptualism, 68
Condemnation, 600
Condensed matter physics, 354
Conditional imperative, 524, 545
Conditional probability, 51

Consequentialism, 550, 558
Contingent, 431
Coordinate geometry, 181
Coordinate mapping, 182
Coordinate systems, 181
Copenhagen interpretation, 315
Copernicus, 153, 178
Coriolis force, 247
Corpuscles, 169, 204
Corruption, 329, 369
Cosmic microwave background, 245
Cosmological argument, 399, 425, 437
Cosmos, 82
Coulomb's law, 210
Covariant derivative, 246, 679
Create, 482
Creation operator, 329, 358, 364, 378, 626, 636, 648
Curl, 213
Curvature, 248
Curved space-time, 244, 380, 678

Darwin, 163
de Soto, 145, 147
Deed ethics, 496
Defect, 573
Defence, 581
Degrees of goodness, 570
Deism, 86, 136, 258
Derived existence, 430
Descartes, 90, 147, 169, 180
Design, 199
Determinism, 85, 164, 165, 224, 264, 326, 668
Differential equation, 202
Differential geometry, 244
Differentiation, 201
Diffraction, 215, 216
Dirac δ-function, 212, 308, 316, 333, 379, 647
Dirac equation, 324, 329
Disorder, 573
Divergence, 212
Diversity, 587
Divine command ethics, 546
Divine response to evil, 618
Divine simplicity, 449
Dualism, 108
Duty ethics, 545
Dynamic block theory of time, 240

Economics, 587
Effective field theory, 354
Effective Hamiltonian, 355
Efficient cause, 115, 125, 258, 264, 267–269, 271, 275, 279, 371, 430, 438, 540, 572
Eigenstate, 312, 367, 377
Eigenvalue, 296
Eigenvector, 296
Einstein, 182, 218, 233, 243, 248, 256
Einstein's field equation, 660, 681
Einstein-Hilbert action, 251, 680
Electric charge, 210, 242
Electric potential, 213, 218
Electro-weak theory, 347
Electromagnetic Lagrangian, 218
Electromagnetism, 210, 237, 323, 333
Electroweak theory, 21, 350
Element, 483
Empiricism, 74, 105, 136, 185, 372
Energy, 204, 210, 236, 242, 257, 266, 292, 313, 380, 653, 668, 676, 677
Energy bands, 355
Energy-momentum, 257
Enlightenment, 14, 282, 395
Epistemology, 373
Equality, 514
Essence, 99
Essential, 99
Essentially ordered series, 402, 403, 406
Eternal, 457
Ethical implications of Schrödinger's cats, 555
Ethics, 497
Ethics from authority, 504
Euclidean geometry, 221, 231
Eudaimonia, 538
Euthyphro argument, 504, 548
Event, 115, 224
Event causality, 24, 119, 265, 326, 362, 371, 654
Everett interpretation, 365, 669, 699
Evil, 500, 540, 574, 595, 609
Evil God objection, 622
Exclusion principle, 322, 331
Existence, 112, 424, 429
Experimental error, 55, 309
Experimental physics, 151, 193, 198, 319

Faith, 593
Falsification, 54, 278
Faraday, 213
Fermat, 207
Fermi's golden rule, 383
Fermion, 322, 331, 637, 647
Feynman, 341, 677
Feynman diagrams, 341, 385, 661
Feynman rules, 342
Field, 211, 213, 220, 326, 353, 638, 658
Final causality, 23
Final Cause, 142
Final cause, 115, 165, 259, 264, 266, 279, 305, 370, 410, 418, 458, 527, 559, 561
Fine tuning, 689
Five element theory, 128
Five ways, 398
Fixed distribution, 300
Fixed points, 297
Fock state, 328, 330, 365
Force, 196, 204, 205
Form, 59, 92, 94, 264, 357, 369, 421, 532, 559, 564
Formal cause, 115, 165
Fortitude, 591
Four vector, 236
Free will, 103, 139, 262, 464, 583, 609, 611, 625, 630, 678, 685, 689
Freedom, 104, 584
Frequentest interpretation of probability, 45
Function, 222

Gödel, 39
Gödel's theorem, 39, 186, 392
Galilean relativity, 225
Galileo, 129, 145, 147, 149, 163, 191, 287
Gauge field, 329, 338, 345, 347, 658, 662
Gauge symmetry, 218, 329, 658
Gauss, 212, 231
General providence, 701
General relativity, 243, 380, 678
Generation, 329, 369
Geodesic, 248, 681
Geometrical physics, 13
Geometrical series, 340
Gilbert, 163, 210
Global symmetry, 660
Gluon, 347, 676
God, 394, 444
God of the gaps, 416
Golden mean, 543
Good, 540, 565
Goodness, 306, 519, 537
Goodness of God, 571, 593
Grace, 600, 601
Gravity, 205, 243, 247, 379, 678
Greed, 589
Greek science, 134
Grosseteste, 154
Growing block theory of time, 239

Hamiltonian, 328, 365, 377, 654
Hamiltonian mechanics, 209, 328
Harm, 502
Harmonic oscillator, 318
Hate, 594
Heisenberg, 279, 367
Heraclitus, 90, 241
Hermitian matrix, 296
Hermitian operator, 311, 316, 642
Heytesbury, 145
Higgs Boson, 350, 671
Hobbes, 165, 169, 185, 507
Hope, 592
Hume, 91, 106, 189, 268, 376, 395, 413, 496, 516, 524, 568, 660, 705, 721
Huygen's principle, 216
Hydrogen atom, 317
Hyperbolic functions, 232
Hyperbolic Geometry, 231
Hyperbolic geometry, 231, 329, 627, 655
Hypostasis, 483

Ideal, 420
Idealism, 68, 185
Imaginary Numbers, 308
Imagination, 105
Immaterial, 97, 454
Immutable, 446
Impartial, 628
Impetus, 130, 146
Implicitly present, 472
Indeterminacy, 104
Indeterminate, 19
Indeterminism, 329, 361, 630, 668, 719
Inertia, 149, 169, 204, 289, 404
Inertial frame, 225
Instrumental cause, 122, 271, 403
Instrumental purpose, 563
Instrumental series, 402, 403, 406, 438
Instrumentalism, 137
Integration, 201
Intellect, 102, 583, 625, 629
Intellectual purpose, 562
Interference, 20, 51, 207, 215, 216, 233, 321, 644, 667
Interpretations of probability, 44
Intuitive ethics, 495
Invariance, 239
Irreducible, 44
Is/ought dichotomy, 516, 524
Islam, 12

John Philoponus, 130
Justice, 581, 591

Kant, 91, 112, 165, 185, 189, 282, 372, 376, 395, 419, 440, 519, 545, 653
Kenny, 428
Kepler, 179
Klein paradox, 325
Klein-Gordon equation, 314, 323, 329
Knowledge, 473
Kolmogorov's axioms, 44

Lagrangian mechanics, 207, 218, 241, 251, 328
Lagrangian physics, 336
Least time principle, 207
Liberal theology, 263
Libertarian ethics, 512
Life, 109, 577
Light waves, 217
Likelihood, 50, 51, 310, 338, 633, 636
Linear operators, 296
Local symmetry, 661
Locality, 330, 338, 362, 652, 653, 669
Locke, 105, 507
Logic, 37, 229
Logical interpretation of probability, 48
Logical problem of evil, 602
Logical system, 39
Logically necessary, 432
Logically necessary being, 423
Lorentz force law, 214
Lorentz symmetry, 234, 236, 328, 333, 654, 657
Love, 590
Lowering operator, 318

Mach's principle, 245
Magnetic potential, 218
Man, 576
Mapping, 222
Market, 589
Markov chains, 300
Mass, 204, 236, 329, 673
Material, 97
Material cause, 115, 264
Mathematical physics, 142, 149, 187, 193, 198
Matrices, 181

Matter, 93, 94, 369
Maxwell, 214, 218, 228, 232, 293, 336, 658
Maxwell's equations, 214, 237
Measurement, 638
Measurement operator, 308
Measurement problem, 315, 362, 366
Measuring goodness, 554, 611
Mechanism, 85, 136, 148, 163, 170, 220, 238, 249, 290, 365, 372, 376
Mechanism, premises, 164
Medieval physics, 139
Medieval universities, 143, 144, 398
Memory, 105
Merton calculators, 144, 371
Metaphysics, 9, 61, 253, 360, 375
Methodological naturalism, 259
Metric tensor, 231, 679
Metrical structure, 231
Michelson-Morely experiment, 233
Mill, 550
Mind, 102, 481
Mind-body problem, 185
Minkowski, 231
Minkowski metric, 236
Miracle, 702
Miracles in other religions, 722
Models, 39
Modernism, 17
Momentum, 204, 210, 236, 242, 257, 266, 313, 315, 328, 653
Money, 587
Moore, 496, 519, 527
Moral good, 567
Moral nihilism, 523, 530
Moral relativism, 515
Motion, 90, 114, 125, 346, 369, 400
Moving spotlight theory of time, 239
Multiverse, 698, 699

Natural, 100
Natural law ethics, 559
Natural motion, 114, 128
Natural purpose, 527, 539, 560, 561
Natural selection, 301
Naturalistic fallacy, 519
Necessary, 431
Neutrino, 347, 670
Neutrino masses, 351
Newton, 163, 197, 218
Newton's laws of motion, 203
Nicolas of Cusa, 155
Noether, 241, 248
Nominalism, 68, 107, 164, 363
Number operator, 309, 314, 329

Objective categorisation, 78
Objective ethics, 500
Observable, 638
Ockham, 144, 397, 512
Omnipotence, 467, 609, 701
Omnipresence, 457
Omniscience, 474, 476, 627
One God More argument, 445
Ontological argument, 391, 412, 422, 440
Operator, 181, 307, 308
Optics, 215
Oresme, 146
Original sin, 595, 613
Orthogonal, 44, 104, 636

Paine, 509
Parmenides, 90, 241
Part, 448
Partial differentiation, 202
Particle, 66, 220, 353
Particle interaction, 339
Path integral, 641, 662, 677
Path ordering, 648
Perception, 104
Person, 483
Perturbation theory, 203, 234, 338
Peter the Pilgrim, 150
Philosophical relativism, 227
Philosophy of physics, 82
Philosophy of time, 239
Phonons, 355
Photon, 333, 336, 658, 674
Physical determinism, 164
Physical law, 82, 462
Pilot wave interpretation, 315, 361, 669
Planck's constant, 312, 317
Plato, 10, 531
Point of view invariance, 659
Positivism, 107, 253
Potentia, 92, 114, 327, 367, 400
Potentiality, 113
Practical ethics, 496
Precision, 78, 138
Predicate, 420
Predicates, 361, 366
Presentism, 224, 239
Primary quantities, 194
Principle of causality, 122, 265, 316, 369, 377, 400
Principle of proportionate causality, 124
Principle of sufficient reason, 65, 122
Probabilistic problem of evil, 615
Probability, 43, 310, 338, 632, 636
Problem of evil, 495, 602, 607
Problem of suffering, 603
Procession, 485
Propagator, 340, 664
Properties, 79, 95, 370
Properties of materials, 355
Property realism, 653
Prudence, 591
Ptolemy, 129, 177, 192
Pure actuality, 401, 448
Purpose, 118
Pursuit of good, 569

Quantum chromodynamics, 21, 99, 347, 670
Quantum electrodynamics, 21, 337, 654
Quantum field theory, 21, 66, 326, 361, 376, 625
Quantum gravity, 251, 381, 678
Quantum mechanics, 313, 323, 360

Quantum physics, 18, 135, 313, 326, 634
Quantum vacuum, 327
Quark, 347, 670

Radioactive decay, 381, 408
Raising operator, 318
Rational agent, 109
Rational animal, 111
Rational soul, 111
Rational universe, 166
Rationalism, 75
Rationality, 64
Realism, 68, 98, 363
Reductionism, 185
Reflection, 215
Reformation, 192
Refraction, 215
Relation, 484
Relative motion, 225
Relativity, 16
Renaissance, 148
Renormalisation, 344, 665
Responsibilities, 509
Resurrection, 609, 615
Ricci scalar, 248
Ricci tensor, 248
Riemann, 244, 627
Riemann curvature tensor, 248, 679
Rights based ethics, 506
Rotation, 232
Russell, 414

Scalar product, 210
Scale invariance, 657
Scepticism, 75
Schleiermacher, 388
Scholastic metaphysics, 93, 285, 367, 390
Schrödinger equation, 314
Science, 133
Scientific method, 53, 255
Scientific realism, 138
Scientific revolution, 114
Scotus, 144
Second law of thermodynamics, 298, 438
Secondary quantities, 194
Secular ethics, 528
Self-awareness, 105
Self-existent, 430
Self-ownership, 507
Set theory, 229
Sign, 703
Simple, 109, 449
Slavery, 619
Social contract, 507, 512
Society, 586
Soul, 110
Space, 81, 222, 256
Space-time, 257, 626
Special providence, 701
Special relativity, 233, 323, 329, 654
Species, 98
Spin, 219, 322, 324, 333
Spinor, 324, 333, 650, 657
Spinoza, 718, 719
Spirit, 111
Spontaneous symmetry breaking, 673
Standard model, 21, 347, 351, 669
State, 64, 92, 295, 307, 628
Statement, 76
Static block theory of time, 240
Statistical equilibrium, 292, 300
Statistical mechanics, 291
Stress-energy tensor, 248, 660, 679
Subjective theory of economic value, 587
Substance, 97, 356
Substance causality, 24, 119, 264, 327, 362, 371, 654
Substance realism, 653
Substantial change, 93
Suffering, 582, 603
Superconductivity, 356
Supernatural, 101
Superposition, 370
Supersymmetry, 690
Syllogism, 40
Symbolic logic, 42
Symmetry, 45, 54, 66, 209, 218, 239, 241, 249, 329, 337, 364, 405, 655, 657, 665, 683

Taylor series, 234, 311
Teleological argument, 410, 458
Teleology, extrinsic, 118
Teleology, intrinsic, 118
Temperance, 591
Temperature, 293
Theism, 25, 141
Theoretical error, 55
Time, 81, 223, 256
Time evolution operator, 308, 641
Time ordering operator, 311, 661
Timeless, 454
Topology, 231
Transfer operator, 655
Trinity, 490
Truth, 7, 76
Tycho, 178

Uncertainty principle, 317, 361
Unification of physics, 218
Uniqueness of God, 453
Unitary matrix, 331, 635, 642
Universal ethics, 501
Universality, 363
Universals, 68, 79
Universe, 82
Unnaturalistic fallacy, 523
Utilitarianism, 550, 567

Vacuum pair production, 340, 385
Vector, 181
Vector product, 213
Vectors, 181
Vice, 531
Violence, Biblical, 620
Violent motion, 128
Virtual existence, 93, 113, 327
Virtue, 531

Virtue ethics, 531, 544, 556, 558
Voluntarism, 515

Wave frequency, 216
Wavefunction collapse, 315
Waves, 215, 313
Wealth, 588
Who created God? argument, 441
Will, 103, 125, 460

If God exists, what sort of universe would He create?

The ancient Greeks thought that the study of physics required listing the possible causes and tendencies of every different type of being. Theists believe that physics is a description of how God sustains the universe. Geometers believe that physics is a mathematical representation of reality; or perhaps a mathematical representation of our opinions about reality. Mechanists say that it is a description of fixed and unbreakable laws, independent of God, which matter is blindly compelled to obey. Empiricists say that it is a collection of experimental data.

Ideas have consequences, and fundamental axioms no less than anything else. It ought to be possible to begin with what we believe physics represents, and through careful reasoning conclude with a picture of its laws. If physics is a description of how God sustains matter, then the laws of physics should in some way reflect God's nature. So what can we deduce concerning the mathematical laws, which today's physicists are painstakingly uncovering, from what theists believe about God? And how well do those predictions line up against reality?

And what about the philosophies which inspired Galileo and Newton, and led to modern atheism? What form would we expect physical law to take if we assume those premises are true? And how well does that line up against our current knowledge, now that Newton's theory has been superseded by quantum field theory and general relativity? Do the renaissance and enlightenment arguments against theism and classical philosophy still stand in the light of today's best science?

At the start of the twentieth century, it seemed as though modern science and theism were in the deepest conflict, from which there could be only one victor. But today, does physics still fight against theism, or has it instead turned its weaponry against atheism?

Cover Image: The Carina Nebula, taken from the Hubble Space Telescope.
Courtesy of: NASA, ESA, N. Smith (U. California, Berkeley) et al., and The Hubble Heritage Team (STScI/AURA)

US:	$31.99
UK:	£23.99
Europe:	€26.99

ISBN 9798568549284